Journal of Business and Psychology
https://doi.org/10.1007/s10869-024-09948-5

ORIGINAL PAPER

Revisiting the Inhibitory Effect of General Mental Ability on Counterproductive Work Behavior: The Case for GMA-Personality Interaction

Alexandra M. Harris-Watson[1] · Joshua D. Miller[2] · Nathan T. Carter[3]

Accepted: 3 April 2024

Abstract
Counterproductive work behavior (CWB) is an important component of job performance that is known to be related to critical personal and organizational consequences. Thus, both researchers and practitioners are interested in better understanding CWB's primary drivers. Despite its popularity, the theoretical *inhibitory effect* of GMA on CWB, which predicts that employees with higher GMA will show lower CWB, has seen weak and inconsistent empirical support. Here, we propose that a reason for this divide between theory and empirical studies can be explained by a more appropriate interpretation of the inhibitory effect as *conditional*, in that the strength of the GMA-CWB relationship is dependent on other critical individual differences. We suggest that the meta-trait *stability*–which subsumes conscientiousness, agreeableness, and emotional stability, the three personality traits shown to be consistently positively related to CWB–is critical for revealing the GMA-CWB relation in empirical studies. Specifically, we hypothesize that the inhibitory effect is dependent on the meta-trait stability such that the expected negative GMA-CWB relationship is strongest for those with low levels of stability but is not apparent for those with high levels of stability. Results supported the conditional inhibitory hypothesis across two large samples. Implications for theory and practice are discussed.

Keywords Intelligence · Personality · Stability · Counterproductive work behavior · Interaction

Inhibitory Effect of General Mental Ability on Counterproductive Work Behavior: The Case for an Interaction

During the last two decades, one major shift in research on selection systems has been an expansion of the criterion domain to include contextual performance in addition to more traditionally emphasized task performance (Borman & Motowidlo, 1993; Sackett & Lievens, 2008). Within contextual performance, counterproductive work behavior (CWB) describes voluntary behavior that harms an organization or the individuals within the organization. Thus, CWB can have serious consequences for both individuals and organizations. For example, theft—one type of CWB—is a large financial risk for organizations (Bennett & Robinson, 2000). Perceptions of CWB also contribute to overall ratings of an individual's job performance (Choi et al., 2018; Lievens et al., 2008; Rotundo & Sackett, 2002). Moreover, CWB can reduce the performance of entire business units, not just the individuals committing CWBs (Dunlop & Lee, 2004). Research also suggests that CWBs can lead to increased experiences of interpersonal incivility and organizational constraints such that CWB leads to more CWB (Meier & Spector, 2013). Consequently, preventing and reducing CWB is of great interest to organizations.

The growing interest in CWB as a criterion for selection systems has invigorated research on its antecedents more broadly. Historically, studies regarding antecedents of CWB have focused on attitudinal and personality variables (Spector & Fox, 2002), perhaps stemming from the assertion by

Additional supplementary materials may be found here by searching on article title https://osf.io/collections/jbp/discover.

Alexandra M. Harris-Watson
amharriswatson@ou.edu

1 Department of Psychology, University of Oklahoma, 455 W. Lindsey Street, Norman, OK 73019, USA

2 University of Georgia, Department of Psychology, Athens, USA

3 Michigan State University, Department of Psychology, East Lansing, USA

Borman and Motowidlo (1993) that personality or "will-do" characteristics should predict contextual performance more so than ability or "can-do" characteristics. However, researchers have increasingly broadened their consideration of other individual differences, particularly in relation to CWB. Indeed, Dilchert et al. (2007)–based on evidence that GMA is negatively related to deviance in the more general sense (Gottfredson & Hirschi, 1990; Lynam et al., 1993)–hypothesized that GMA *should* inhibit (i.e., be negatively related to) CWB, termed the *inhibitory effect*. Despite its theoretical appeal and popularity (e.g., see Ones et al., 2012), empirical support for this hypothesis has been mixed at best. Early studies showed some support for GMA as a negative predictor of CWB (Dilchert et al., 2007), but a subsequent meta-analysis found no support for a GMA-CWB relationship (Gonzalez-Mulé et al., 2014). Thus, empirical findings seem to conflict with researchers' general agreement that GMA should meaningfully predict CWB.

In this article, we offer an explanation that focuses on redefining the popular inhibitory effect hypothesis (Dilchert et al., 2007) as a *conditional* inhibitory effect (i.e., an interaction). Specifically, we suggest the traditionally hypothesized negative influence of GMA on CWB emerges only among individuals low in other dispositional characteristics that decrease their likelihood of engaging in CWB. We draw on a robust body of literature supporting a relationship between stability–a higher-order "meta-trait" comprised of the FFM traits conscientiousness, extraversion, and emotional stability–and CWB (Berry et al., 2007; DeYoung et al., 2008) to suggest there is an *interactive* effect of GMA and stability on CWB, such that the inhibitory effect is conditional on levels of stability.

The Inhibitory Effect: The Case for a GMA-CWB Relationship

The primary theoretical rationale for a negative relationship between GMA and CWB draws on evidence for a negative GMA-deviance relationship in the criminology literature. Prior scholars have suggested that CWBs show substantial overlap with deviant behavior generally. Some behaviors classified as CWBs are in fact illegal, such as theft, white collar crime, violence, or use of illicit drugs on the job (Bowling & Gruys, 2010; Ones, 2002), and those CWBs that are not criminal still "involve breaking rules and deviating from established norms" (Dilchert et al., 2007). Because CWBs reflect rule-breaking and norm-violation similar to deviant behavior, the GMA-deviance link might help to explain CWB (Dilchert et al., 2007; Ones et al., 2012). Notably, prior work has shown that the strongest trait predictors of criminal behavior are also important predictors of CWB (Marcus & Schuler, 2004), giving some support to the idea that predictors of deviance broadly may also predict CWB.

In particular, there are two popular theoretical rationales for a GMA-deviance relationship that researchers have extended to postulate a GMA-CWB relationship: (a) an *inhibitory effect*; and (b) *moral reasoning* (Dilchert et al., 2007; Gonzalez-Mulé et al., 2014; Ones et al., 2012). The inhibitory effect suggests that because high-GMA individuals possess greater capacity to evaluate the long-term consequences of their actions, high-GMA individuals will be more likely to recognize the negative consequences of CWBs and adjust their behavior accordingly. Relatedly, the moral reasoning justification for a GMA-deviance link suggests that GMA *indirectly* affects deviance via individuals' increased capacity to evaluate the "rightness" or "wrongness" of their actions. That is, the moral reasoning rationale likewise suggests that GMA should inhibit CWB but that the relationship is mediated through the related construct of moral reasoning (Dilchert et al., 2007). Whereas the inhibitory effect broadly encompasses a variety of reasons that GMA would suppress CWBs, the moral reasoning rationale specifies that the inhibitory effect happens via moral reasoning. Effectively, the moral reasoning rationale is a narrower specification of the inhibitory effect.

Although both the inhibitory effect and moral reasoning rationales are drawn from the criminology literature, empirical support has only been shown for the broader inhibitory effect. There is substantial empirical evidence for a negative relationship between GMA and delinquency (Gottfredson & Hirschi, 1990), and this negative relationship persists even after controlling for other contextual factors such as socioeconomic status (Lynam et al., 1993). In contrast, research findings have so far not shown much support for the type of mediation pathway predicted by the moral reasoning rationale (Lykken, 1991; cf Dilchert et al., 2007). Thus, while empirical evidence for the inhibitory effect is well-established within the criminology literature, consistent empirical evidence for the narrower moral reasoning rationale has not yet been shown. Given limited support for moral reasoning as a mediator within the criminology literature and theoretical overlap between the moral reasoning and inhibitory effect rationales, we believe the inhibitory effect offers the most compelling theoretical framework from the criminology literature that might be extended to the workplace deviance literature.

Despite ample support for an inhibitory effect within the criminology literature and the hypothesis that GMA should likewise predict CWB, empirical evidence has not universally supported that relationship. The earliest direct test of the inhibitory effect on CWB found that GMA and CWB show a modest, negative relationship (Dilchert et al., 2007). Moreover, in their review of cognitive ability in personnel selection, Ones et al. (2012) summarized results of earlier atheoretical, large-scale validation efforts (McHenry et al., 1990; Oppler et al., 2001) and concluded that results showed support for an inhibitory effect of GMA on CWB. However, in a more recent meta-analysis, Gonzalez-Mulé et al. (2014) found no evidence of a clear relationship between GMA and CWB, which suggests no inhibitory effect of GMA on CWB. Thus, empirical findings regarding the GMA-CWB relationship have been inconsistent and do not clearly support an inhibitory effect of GMA on CWB.

The Conditional Inhibitory Effect: The Case for an Interaction

One possible explanation for the lack of support for an inhibitory effect is that the inhibitory effect has thus far been operationalized inconsistently with its theoretical form. To date, the inhibitory effect of GMA has been operationalized primarily as a *main* effect of GMA on CWB. We argue that a theoretical inhibitory relationship between GMA and CWB instead reflects a *conditional* effect of GMA (i.e., an interaction) such that GMA affects CWB only in the presence of other CWB-risk factors. Returning to the key mechanism of the inhibitory effect, this theory suggests that high GMA enables individuals to better anticipate the consequences of their actions and adjust behavior accordingly. However, inhibition assumes individuals have some proclivity to engage in CWB in the first place. Without such an initial inclination to engage in CWB, the benefits of GMA for suppressing CWB are not needed.

Indeed, other researchers have similarly suggested that the inhibitory effect may be most pronounced in the context of other deviance risk factors. Dilchert et al. (2007) suggest:

> . . . the investigation of potential interaction effects could further contribute to researchers' understanding of the observed relationships. For example, it seems conceivable that the negative cognitive ability-CWB association would be even stronger among individuals who lack other personal characteristics that act as preventative factors. (p. 624)

That is, the inhibitory effect of GMA may be strongest among individuals with "other personal characteristics" that make them prone to engaging in CWB. Similarly, Gonzalez-Mulé et al. (2014) allude to an interaction between GMA and other risk factors, suggesting high-GMA individuals have "... less likelihood of falling victim" to "the vicious frustration-aggression cycle" (p. 1224). The "frustration-aggression" cycle refers to evidence that other, non-GMA antecedents affect CWB and that GMA may weaken those relationships.

This type of conditional effect in which the relationship between GMA and CWB is dependent on other variables reflects an interaction. Thus, whereas prior research has framed the inhibitory effect as a simple direct effect, we argue that inhibition is most appropriately framed as a *conditional* or *interactive* effect. GMA may inhibit CWB only when individuals possess some other tendency to engage in CWB. In contrast, individuals who possess no other risk factors for CWB would be unlikely to engage in CWB regardless of GMA, and GMA would show no effect on CWB.

Notably, the two quotes above–though both referencing an interaction–use language that refers to slightly different framings of an interaction. The first uses language that describes GMA as *moderated by* other risk factors (i.e., the negative GMA-CWB relationship is strongest at certain levels of other characteristics), whereas the second refers to GMA as *moderating* other risk factors (i.e., GMA weakens the relationship between other characteristics and CWB). Although, statistically, an interaction between two variables is equivalent regardless of which variable is treated as the moderator, here we focus on the framing of GMA as the independent variable and other dispositional tendences as the moderator. This framing is most consistent with our theory that inhibitory effect of GMA on CWB is *dependent on* other risk factors. That is, we hypothesize an inhibitory, negative GMA-CWB relationship but only among those with levels of other characteristics that predispose them to CWB.

Moreover, because prior work on the inhibitory effect assumes that GMA shows a simple, direct effect on GMA, framing the interaction in this way allows for more direct comparisons with prior findings. For example, we calculate Johnson-Neyman (JN) confidence intervals (CIs) to determine at what levels of the moderator the effect of GMA on CWB is significant. We also graph the influence of GMA on CWB at differing levels of the moderator (as opposed to the influence of other characteristics at varying levels of GMA) to illustrate how the GMA-CWB relationship changes dependent on the moderator.

The Meta-trait of Stability

Interpreting the inhibitory effect as conditional requires considering what other variables are critical risk factors for CWB. Although literature on the relationship between GMA and CWB is somewhat limited, there is a robust body of literature on the importance of personality in predicting CWB. Much of the research on the relationship between personality and CWB has focused on personality as it relates to control (Marcus & Schuler, 2004; Spector, 2011), in part due to the popular theory that self-control is the most important predictor of deviance generally (Gottfredson & Hirschi, 1990; Pratt & Cullen, 2000; Vazsonyi et al., 2001). This theory suggests that self-control is critical in preventing deviant behavior because nearly all such behaviors involve "immediate and easy gratification of desires at the cost of possible long-term negative consequences" (Hirschi & Gottfredson, 1994; Marcus & Schuler, 2004, p. 648). Marcus and Schuler (2004) extend this theory to general counterproductive behavior at work and show that, even after accounting for many other individual differences, self-control is strongly negatively related to CWB.

A closely related concept to self-control is the meta-trait *stability,* a higher-order dimension comprised of the Big Five traits conscientiousness, agreeableness, and emotional stability. Although the Big Five were originally thought to be orthogonal, they in fact co-vary and load onto two higher-order dimensions, or meta-traits (Digman, 1997; DeYoung, 2015; DeYoung, 2006). Because several researchers independently derived the meta-traits, they have been given alternative labels at different times. Here, we use the label "stability" to refer to the meta-trait comprised of conscientiousness, agreeableness, and stability, which is consistent with the terminology proposed by DeYoung (2006). Notably, the meta-trait now known as stability has also been labelled *self-control* (Olson, 2005) and *social self-regulation* (Saucier et al., 2014), reflecting the centrality of behavioral restraint and regulation to the trait.[1]

In his Cybernetic Big Five theory, DeYoung (2015) attempts to reconcile empirical support for the meta-traits with a theory of their emergence and function. DeYoung suggests that overall stability promotes the "protection of goals, interpretations, and strategies, from disruption by impulses" (p. 10), and each of the three traits plays a role in preventing such impulses. Perhaps most obviously, emotional stability reflects stability of mood and emotion. Conscientiousness reflects "motivational stability, maintaining progress toward long-term or abstract goals" and agreeableness reflects "social stability, maintaining the harmony of social interactions" (DeYoung, 2015, p. 47). Thus, emotional stability, conscientiousness, and agreeableness protect against disruption in emotional, motivational, and social impulses, respectively. Indeed, prior research has shown that measures of self-control are strongly related to conscientiousness, agreeableness, and emotional stability (de Vries & van Gelder, 2013; Marcus, 2003; O'Gorman & Baxter, 2002).

The role of stability in constraining behavior may stem from its neurobiological underpinnings. DeYoung and colleagues (2010, 2015) suggest that the meta-traits may emerge in part from individual differences in neurotransmitter functioning. Specifically, stability is thought to be related to serotonin functioning such that low levels of serotonin are positively related to impulsiveness and aggression (Carver & Miller, 2006; Hirsh et al., 2009). Serotonin may help facilitate long-term goal prioritization by signaling danger, as well as suppressing "negative affect and aggression while maintaining behavior and motivational stability" to avoid or overcome the danger (Hirsh et al., 2009, p. 1086). Serotonergic functioning is positively related to all three traits comprising stability (conscientiousness, agreeableness, and emotional stability). Thus, their shared variance (i.e., the overlap in all three with stability) may be attributable to serotonergic functioning and its effect on behavioral constraint.

Although we are not aware of any empirical work that examines the effect of stability on CWB, related research strongly suggests that stability is a critical predictor of CWB. First, stability as a construct is consistent with self-control which, as outlined above, is thought to be an important—if not the *single* most important—predictor of deviance generally and CWB. Second, stability has been shown to predict related constructs. Stability is positively related to conformity with social norms (DeYoung et al., 2002) and negatively related to externalizing behavior broadly (DeYoung et al., 2008). Finally, research on the relationships between the Big Five traits and CWB clearly supports a negative relationship between CWB and all three of the traits that comprise stability (conscientiousness, agreeableness, and emotional stability; Berry et al., 2007). Thus, we would also expect stability to negatively predict CWB.

To the extent that GMA exhibits a negative relationship with CWB only when individuals possess other critical risk factors, *low stability* seems likely to be one such factor. Thus, we expect an interaction between GMA and the meta-trait stability (see Fig. 1) in which the GMA-CWB relationship is conditioned on (i.e., moderated by) stability. Because low stability reflects an individual's general tendency toward impulsive behavior and immediate gratification, high GMA may help an individual to better understand the consequences of such behavior and curb impulses accordingly. However, high GMA is not as likely to benefit high-stability individuals who are not prone to engaging in CWB.

Hypothesis 1: The meta-trait stability moderates the GMA-CWB relationship such that GMA shows a negative relationship with CWB among low-stability, but not high-stability, individuals.

In the current study, we consider only the influence of a GMA-stability interaction on CWB. However, theoretically, the interpretation of the inhibitory effect of GMA as conditional also applies to deviance generally. That is, we would

[1] The second meta-trait *plasticity* encompasses the remaining Big Five traits of openness and extraversion and concerns the creation and exploration of new goals. Importantly, the two meta-traits do *not* reflect opposing ends of the same dimension. As noted by DeYoung (2015), "The opposite of stability is not plasticity but instability, and the opposite of plasticity is not stability but rigidity" (p. 47). See DeYoung et al., (2002, 2010) for further reading on possible neurobiological substrates of the Big Five traits, including the relationship between plasticity and dopaminergic functioning.

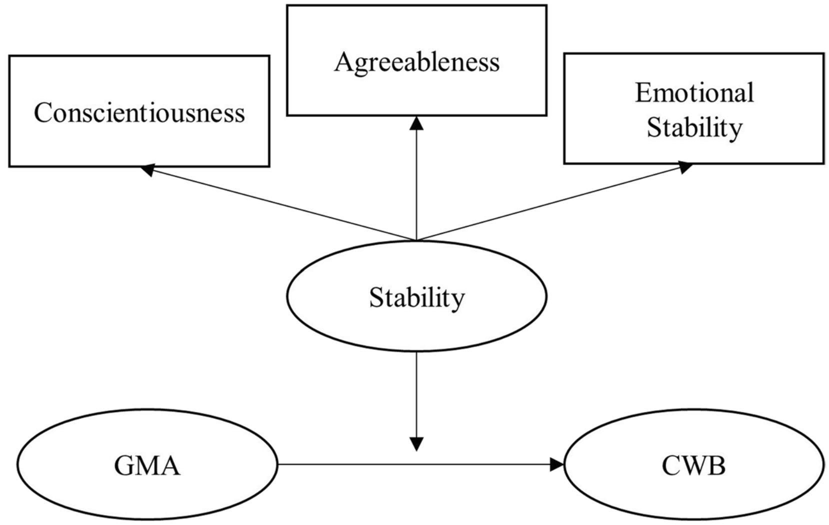

Fig. 1 Expected relationship between GMA and CWB moderated by meta-trait stability (i.e., composite of Conscientiousness, Agreeableness, and Emotional Stability)

expect GMA to have the strongest inhibitory effect on deviance among those who are otherwise predisposed to engage in deviant behavior (i.e., low-stability individuals). Although prior research clearly supports a significant main effect of GMA on deviance generally, this is not evidence *against* an interaction. Rather, statistically, main effects reflect the influence of the predictor on the criterion *across* the range of moderator values in a sample (Busenbark et al., 2022). If the effect of GMA on deviance is negative at low levels of stability and effectively zero at high levels of stability, the "main" effect of GMA on deviance without accounting for stability would often appear to be negative (i.e., "averaged" across levels of stability).

Support for a negative GMA-deviance but *not* a GMA-CWB link to date may be attributable to mean differences in stability across sampled populations. Because stability-related traits show relationships with other important career and work outcomes (Barrick et al., 2001; Barrick & Mount, 1991; Mount et al., 2006), many workplaces select based on stability-related traits. Such selection may happen both directly through selection assessment and indirectly through employment history (e.g., very low stability individuals may be less likely to maintain long-term employment). In contrast, research on the inhibitory effect in criminology most commonly focuses on adolescents. Not only does self-control increase throughout adolescence, but self-control in adolescence is related to work outcomes among adults (Allemand et al., 2019). Thus, we would expect samples of working adults to have higher mean-level stability than broad samples of adolescents.

These mean differences in samples may explain why, to date, evidence supports a main effect of GMA on deviance but not on CWB. If the inhibitory or *negative* effect of GMA on deviant behavior is in fact dependent on low levels of stability as hypothesized, an independent main effect of GMA may appear more consistently in samples in which stability tends to be lower (e.g., adolescents) relative to samples in which stability is slightly higher (e.g., working adults). Thus, a main effect of GMA on deviance generally is not theoretically inconsistent with the broader interpretation of the inhibitory effect as conditioned on stability. In fact, the moderating effect of stability may clarify why, to date, evidence supports a main effect of GMA on deviance generally but not CWB.

The Current Study

In the current study, we investigate the possible inhibitory, moderating influence of stability on the relationship between GMA and CWB. Importantly, a variety of other explanations have been proposed for why empirical evidence to date may not support a GMA-CWB relationship. These explanations do not necessarily conflict with the inhibitory interaction proposed in the current study. Thus, we address some of these concerns here.

One suggestion for why GMA and CWB have not shown the expected negative relationship in prior research is that relationship may vary depending on the sub-dimension of CWB (Gonzalez-Mulé et al., 2014). Although a wide variety of CWB dimensions have been considered (Cullen & Sackett, 2004; Spector et al., 2006), a two-dimensional structure that differentiates between *interpersonally* and *organizationally* targeted CWBs (CWB-I and CWB-O, respectively) is generally most popular (Berry et al., 2007; Sackett et al., 2006). Scholars have suggested that differential relationships between GMA and CWB subtypes might obscure a direct relationship between GMA and overall CWB (Gonzalez-Mulé et al., 2014). This same logic also extends to a GMA-stability interaction; differential interactive efffects for the CWB subtypes may obscure an interactive effect on overall CWB. Alternatively, there may also be evidence of GMA-stabiltiy interaction for overall CWB if only one of the the subtypes shows evidence of an interaction but that effect is particularly strong.

Notably, some limited research has already tested related versions of this GMA-stability interaction with mixed support. Differences in findings may be in part attributable to different criteria. Postlethwaite et al. (2009) found that GMA showed the strongest relationship with safety behavior among low-conscientiousness individuals. On the other hand, Ayduk et al. (2007) found that verbal intelligence showed a stronger relationship with aggression among boys with high self-regulation strategies. In this case, aggression is more like CWB-I than CWB-O, whereas safety behavior is more like CWB-O than CWB-I, suggesting that GMA may show differential relationships with CWB depending on the target. Thus, these findings are somewhat consistent with the suggestion by prior researchers that the GMA-CWB relationship varies by CWB subdimension.

Still, research to date is not clear on the theoretical rationale for whether GMA should show a stronger relationship with CWB-I or CWB-O, and there is no empirical evidence for one over the other. Indeed, in their meta-analysis Gonzalez-Mulé et al. (2014) provided rationales for how GMA may inhibit both CWB subtypes, and a GMA-stability interaction may similarly exist for either subtype. For example, a stronger CWB-O relationship may suggest that high-GMA individuals have an easier time anticipating the long-term consequences of CWB-O. Alternatively, a stronger CWB-I relationship may suggest that high-GMA individuals recognize that CWB-I is more readily observed and therefore problematic. In either case, GMA acts to inhibit CWB. Thus, because differential effects by CWB subtype might affect an overall interaction, in addition to considering an interaction between GMA and stability in overall CWB, we also consider such an interaction for CWB-I and CWB-O separately.

> **Research Question 1:** Does the interactive influence of GMA and meta-trait stability differ by CWB subtype (CWB-I and CWB-O)?

Similarly, just as there may be differences in a GMA-stability interaction by CWB subtypes, there may also be differences in how the three lower-order traits (i.e., conscientiousness, agreeableness, and emotional stability) interact with GMA. As with CWB subtype, differential trait-level relationships may obscure an interaction between GMA and the meta-trait stability. Conversely, if one trait shows a particularly strong interaction with GMA, that interaction may mask null effects for the other lower-order traits.

Still, the theoretical argument proposed here for an interaction between GMA and meta-trait stability does not draw a distinction between the three traits. That is, our argument for the moderating influence of stability focuses on the overall functional mechanism of the meta-trait (i.e., prohibition of impulses and prioritization of long-term goals) and not the unique functions of the individual traits. It is not a priori apparent why GMA would differentially interact with the traits in predicting CWB. We expect individuals with low conscientiousness, low agreeableness, and low emotional stability to all benefit from high GMA in curbing CWB.

Moreover, although there is some evidence that conscientiousness and agreeableness show stronger relationships with CWB-O and CWB-I, respectively, both traits still show moderately strong relationships with both CWB subtypes (Berry et al., 2007). Additionally, emotional stability shows similar relationships with both subtypes and, most importantly, all three traits show relationships with overall CWB. Thus, we do not see a meaningful conceptual distinction between these three traits under our outlined theoretical approach and believe that the meta-trait stability is an appropriate construct level for the current theory (Ones & Viswesvaran, 1996).

Nonetheless, if differential effects at the lower-trait level do exist, they are likely to be obscured at the meta-trait level. Further, no prior work has examined the relationship between the meta-trait stability and CWB. Although we believe that there is a strong theoretical rationale and empirical rationale (because all three lower-order traits are negatively related to CWB) for the importance of stability in predicting CWB, it is possible that a moderating effect of stability may be explained by only one or two of these lower-order traits. To evaluate the extent to which the interaction proposed here is driven by the overall meta-trait of stability rather than specific lower-order components, we consider the possibility that the traits comprising stability differentially moderate the GMA-CWB relationship in an exploratory research question.

> **Research Question 2:** Do the three traits comprising the meta-trait stability (i.e., conscientiousness, agreeableness, and emotional stability) moderate the GMA-CWB relationship in the same way as meta-trait stability?

Finally, another popular explanation for the lack of a GMA-CWB relationship is *differential detection* (Marcus et al., 2009). The differential detection hypothesis suggests that a negative GMA-deviance correlation might be expected because higher GMA individuals are less likely to be *detected* when engaging in deviant behavior relative to lower GMA individuals, not because they are less likely to *engage* in deviant behavior (Moffitt & Silva, 1988). This argument suggests that GMA would show a negative relationship with objective and other-rated CWB (e.g., supervisor reports) but not necessarily self-reported CWB.

Table 1 Model-data fit statistics for IRT-scored variables in Sample 1

Variable	Model	M_2	*df*	RMSEA	RMSEA 95% CI Low	RMSEA 95% CI High	SRMSR	TLI	CFI	r_{xx}	% Items RMSEA > 0.05
Conscientiousness (9 items)[a]	GGUM	365.34	18	0.117	0.107	0.128	0.101	0.963	0.976	0.89	0%
Agreeableness[a]	GGUM	1066.58	25	0.172	0.163	0.181	0.218	0.898	0.927	0.84	10%
Emotional Stability[a]	GGUM	610.85	25	0.129	0.120	0.138	0.112	0.930	0.950	0.94	0%
GMA	2PL	225.62	119	0.025	0.010	0.030	0.044	0.981	0.983	0.71	0%
CWB-I[a]	GRM	64.53	14	0.051	0.038	0.063	0.034	0.991	0.994	0.77	0%
CWB-O[a]	GRM	583.23	54	0.083	0.078	0.090	0.067	0.946	0.956	0.83	0%

N = 1,412; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *GGUM* Generalized graded unfolding model; *2PL* Two-parameter logistic model; *GRM* Graded response model; M_2 Limited information goodness-of-fit test statistic; *RMSEA* Root mean square error of approximation; *TLI* Tucker-Lewis Index; *CFI* Comparative fit index; r_{xx} = IRT-based marginal reliability. [a] C_2 statistics used due to low degrees of freedom

To account for the possible influence of differential detection, we use self-reported CWBs only. Self-reported CWB limits the potential confounds of differential detection because it affords a more direct test of whether high-GMA individuals engage in less CWB than do other-reports, which test for differences in detection. Moreover, Berry et al. (2012) found that self-reports are not inferior to other-reports of CWB. In fact, self-reported CWB frequency tends to be slightly higher than other-reported CWB frequency. Further, we utilize samples of employees from Amazon Mechanical Turk (MTurk), as opposed to employees at a specific organization, to increase anonymity safeguards. Whereas research conducted in partnership with specific organizations may lead respondents to believe their data could be shared with a superior, MTurk participants have no reason to believe that their information will be shared with employers.

Thus, we utilize two samples of self-report data from large samples of full-time employees in a variety of occupational contexts via MTurk. Across both samples, we investigate the interactive influence of GMA and stability on CWB, including both CWB-I and CWB-O as mentioned above, and consider possible differential interactions at the trait-level.

Method

Samples

Participants in both samples were recruited through Amazon Mechanical Turk (MTurk). To qualify, participants were required to be working full-time, living in the United States, at least 18 years old, and able to read English. Qualifications were determined based on self-report. Sample 1 data was collected in May 2016 and participants were compensated $0.75 for an approximately 30-min survey. Sample 2 data was collected in March 2020, and compensation was raised to $3.00 for an approximately 30-min survey. Participants were removed prior to analysis for incomplete responses and evidence of inattentive responding (e.g., attention checks, unreasonable response times, and in Sample 2 illogical responses to open-ended questions intended to verify that participants understood English). After this screening, 60% of respondents in Sample 1 were retained for a final sample of 1,412 participants (M_{age} = 34.20, 66% female, 78% white). Sample 1 data have been previously published (see data transparency index for details). In Sample 2, 74% of participants were retained for a final sample of 1,241 participants (M_{age} = 37.65, 46% female, 74% white).

Measures and Scoring

Prior research has demonstrated that using item response theory (IRT)-based estimates in moderated multiple regression reduces Type 1 error rates relative to using sum-scores (Morse et al., 2012). More recently, researchers have also demonstrated that appropriately matching the IRT model to the corresponding response process is essential for reducing Type 1 error (e.g., ideal point versus dominance; Cao et al., 2018; Carter et al., 2014, 2017). Thus, in the current study, all measures are scored using appropriate item response theory (IRT) models. We note the rationale for each model choice below. Tables 1 and 2 display model-data fit statistics for all IRT-scored variables in Sample 1 and 2, respectively. All goodness of fit indices are based on the M_2 statistic, which is analogous to the χ^2 statistic in CFA models but is more appropriate for categorical models such as IRT (Maydeu-Oliveraes & Joe, 2014).

Table 2 Model-data fit statistics for IRT-scored variables in Sample 2

Variable	Model	M_2	*df*	RMSEA	RMSEA 95% CI Low	RMSEA 95% CI High	SRMSR	TLI	CFI	r_{xx}	% Items RMSEA > 0.05
Conscientiousness	GGUM	2622.07	132	0.123	0.119	0.128	0.244	0.938	0.947	0.94	0%
Agreeableness	GGUM	767.01	132	0.062	0.058	0.067	0.129	0.976	0.980	0.97	0%
Emotional stability	GGUM	1137.39	132	0.078	0.074	0.083	0.088	0.960	0.966	0.97	0%
GMA (spatial)	2PL	115.71	35	0.043	0.035	0.051	0.038	0.958	0.967	0.71	0%
GMA (verbal)	2PL	57.95	26	0.032	0.021	0.042	0.034	0.970	0.978	0.62[b]	0%
CWB-I[a]	GRM	72.88	14	0.058	0.045	0.072	0.028	0.993	0.995	0.80	0%
CWB-O[a]	GRM	771.92	54	0.104	0.097	0.110	0.066	0.963	0.970	0.89	0%

N = 1,241; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *GGUM* Generalized graded unfolding model; *2PL* Two-parameter logistic model; *GRM* Graded response model; M_2 Limited information goodness-of-fit test statistic; *RMSEA* Root mean square error of approximation; *TLI* Tucker-Lewis Index; *CFI* Comparative fit index.; r_{xx} = IRT-based marginal reliability

[a] C_2 statistics used due to low degrees of freedom

[b] average reliability across factors

Sample 1

General Mental Ability Cognitive ability was assessed with a subset of the Sandia Matrices items (Matzen et al., 2010), which are free nonverbal matrix-type problems designed similarly to the proprietary Raven's Progressive Matrices (Raven et al., 1998). Respondents are asked to select an image that completes a pattern. Participants completed a subset of items selected to reflect a range of difficulty levels and item-types (Matzen et al., 2010). In their norming study, Matzen et al. (2010) identified a reliability coefficient that exceeded that of the Standard Progressive Matrices on which the Sandia Matrices were based (Sandia Matrices, α = 0.76; Standard Progressive Matrices, α = 0.73).

Answers were coded dichotomously as either correct or incorrect and scored with the two-parameter logistic (2PL) item response theory (IRT) model, which is appropriate for such dichotomously scored items (Embretson & Reise, 2000). After removing items that demonstrated inappropriate fit or item parameters, a final set of 17 items was used to estimate GMA scores. Absolute and relative fit indices suggested excellent model-data fit (RMSEA = 0.025, TLI = 0.98, CFI = 0.98) for the final set of Sandia Matrices items scored with the 2PL.

Personality Conscientiousness, agreeableness, and emotional stability were measured using 10-item scales from the International Personality Item Pool (Goldberg et al., 2006) designed to reflect those in the Revised NEO Personality Inventory (Costa & McCrae, 1992). Respondents indicated the extent to which they agreed that an item describes them accurately on a scale of 1 = "strongly disagree" to 5 = "strongly agree." Consistent with prior research that suggests ideal point IRT models more accurately represent responses to Likert-type personality measures than do dominance measures (LaPalme et al., 2018), the generalized graded unfolding model (GGUM) was used to score all personality scales (Roberts et al., 2000). Notably, due to low degrees of freedom, C_2 fit statistics were calculated (Monroe & Cai, 2015). Although absolute fit indices were high for all measures (RMSEA > 0.08), relative fit indices suggest good model-data fit (TLI & CFI > 0.90) for all three personality measures (Hu & Bentler, 1999). Absolute fit indices may be inappropriate due to low degrees of freedom in the current models and the use of the C_2 statistic. Moreover, marginal reliability was adequate for all three measures. Item-level RMSEA indicated adequate fit for agreeableness and emotional stability but poor fit for the conscientiousness measure (Lee & Luna-Bazaldua, 2019). Consequently, after reviewing item properties, one item with extreme item parameters ("I do just enough to get by") was removed. After removing this item, item-level RMSEA suggested adequate fit of the conscientiousness measure. Thus, we proceeded with a 9- rather than 10-item measure of conscientiousness. An overall stability composite was created by standardizing and then averaging the three IRT-estimated scores for conscientiousness, agreeableness, and emotional stability.

Counterproductive Work Behavior CWB was measured using the popular Workplace Deviance Measure (Bennett & Robinson, 2000), which subdivides CWB into CWB-I and CWB-O. The Workplace Deviance Measure asks respondents to indicate the frequency with which they engage in a variety of behaviors on a scale of 1 = "never" to 7 = "daily." Because the Workplace Deviance Measure is a behavior frequency measure, the graded response model (GRM) was used to score CWB (Samejima, 1997). CWB-I and CWB-O subdimensions were scored separately and then averaged to derive an overall score for CWB.

Sample 2

General Mental Ability Two measures of GMA were administered in Sample 2. As in Sample 1, a subset of Sandia Matrices items was administered (Matzen et al., 2010). Participants completed a 10-item measure designed to represent items with a range of difficulties and appropriate item parameters (Harris et al., 2020). Because the Sandia Matrices are matrix-completion problems and may therefore primarily reflect spatial reasoning, participants also completed a 9-item[2] verbal GMA measure comprised of verbal reasoning as well as letter and number series (i.e., choosing the next letter or number to complete a sequence) item types from the International Cognitive Ability Resource (The International Cognitive Ability Resource Team, 2014). Answers to both the Sandia Matrices and ICAR items were coded dichotomously as either correct or incorrect and scored with the 2PL IRT model. The verbal measure was scored using a 2-factor model in which verbal reasoning and series items comprised separate factors that were allowed to covary. Both absolute and relative fit statistics indicated excellent model-data fit for the Sandia Matrices measure (RMSEA = 0.043, TLI = 0.96, CFI = 0.97) and the final set of ICAR items (RMSEA = 0.032, TLI = 0.97, CFI = 0.98). An overall GMA composite was created by standardizing and then averaging the IRT-estimated scores for both GMA measure types.[3]

Personality Conscientiousness, agreeableness, and emotional stability were measured using the IPIP-120 (Maples et al., 2014), a free measure of the FFM. The IPIP-120 is a 120-item self-report measure similar to the NEO PI-R (Costa & McCrae, 1992) with 24 items per trait. Respondents indicated the extent to which they agree that an item describes them accurately on a scale of 1 = "strongly disagree" to 6 = "strongly agree." As in Sample 1, traits were scored using the GGUM, and an overall stability composite was created by averaging the three scores. Absolute fit indices were again somewhat high, likely due to low degrees of freedom. However, all other fit indices suggested good model-data fit for all three personality measures.

Counterproductive Work Behavior As in Sample 1, CWB was measured using the Workplace Deviance Measure (Bennett & Robinson, 2000). Subdimensions were scored using the GRM and then averaged to create and overall CWB composite.

Data Analysis

Moderated multiple regression models were estimated for each of the criterion variables (overall CWB, CWB-O, and CWB-I) using the lavaan package in R (Rosseel, 2012). The lavaan package allows regression models to be estimated using maximum likelihood with robust standard errors (MLR). MLR estimation was used to account for non-normal distribution of CWB (positive skew is common in deviance research, including CWB; Guay et al., 2016; Penney & Spector, 2005; van Zyl & de Bruin, 2018). First, we present results for our hypothesis concerning the interactive influence of stability and GMA on overall CWB. Consistent with best practices in directional hypothesis testing, reported p-values are one-tailed (Cho & Abe, 2013). Then, we present results regarding a GMA-stability interaction for the two CWB subdimensions, CWB-I and CWB-O (RQ1). Finally, we present results for interactions between the trait-level personality terms (conscientiousness, agreeableness, and emotional stability) and GMA for overall CWB (RQ2). P-values for these exploratory analyses are two-tailed.

Additionally, for each model presented here, we also tested models in which quadratic terms for each variable were included to account for possible collinearity and resultant Type 1 and Type 2 error (Cortina, 1993). In no case did including quadratic terms affect the significance of findings. Thus, only the results of models without such terms are presented for the sake of clarity. Sensitivity analyses were conducted using G*Power (Faul et al., 2009) and suggested that given sample sizes of $N_1 = 1{,}412$ and $N = 1{,}241$, both samples were powered to detect an effect size (ΔR^2 attributable to the interaction) of 0.006.

To illustrate the size of the proposed moderation effect, we provide both a traditional index, ΔR^2, and an alternative effect size index recently proposed by Liu and Yuan (2020), ΔR^2_{mo}. Liu and Yuan argue that ΔR^2_{mo} is a more conceptually appropriate representation of true moderation effect size. Consider a statistical moderation in which the relationship between predictor X and outcome Y varies as a function of moderator Z. Typically, tests of moderation aim to answer the question, "How does the relationship between X and Y vary as a function of Z?" However, Liu and Yuan (2020) suggest that traditional indices such as ΔR^2 do not directly

[2] A 10-item verbal measure was administered. However, to due to an error in administration for one of the items, only 9-items were used to score verbal GMA.

[3] To evaluate whether intelligence type may have affected findings, we also estimated separate models for an interaction between stability and *verbal* intelligence and an interaction between stability and *fluid* intelligence (i.e., spatial measure). Results for fluid and verbal intelligence were not notably different from each other, nor were they different from results using the average of both intelligence types presented in-text.

answer this question. ΔR^2 reflects the variance in *Y* uniquely attributed to the product term *XZ*, divided by the *total variance* of *Y*. Liu and Yuan suggest that a true index of moderation should be based not on total variance but rather "the varying relationship between *X* and *Y*, or the variance of *Y* related to *X* that is further explained by moderator *Z*" (2020, p. 682). The proposed effect size of ΔR^2_{mo} reflects this more appropriate baseline variance, in which *mo* denotes "moderation." Notably, the use of the larger denominator in ΔR^2 (i.e., total variance) is one reason that moderation studies typically report very small effect sizes (Aguinis et al., 2005).

Due to these limitations of ΔR^2 for demonstrating the meaningful impact of moderation effect sizes, we report both ΔR^2 and the newly proposed $\Delta R^2_{mo.}$ In the current study, ΔR^2 represents the proportion of variance in the *CWB criterion* accounted for by the multiplicative GMA-personality term (e.g., GMA by stability, GMA by conscientiousness, etc.). In contrast, ΔR^2_{mo} represents the variability in the *GMA-CWB relationship* that is accounted for by the personality term and thus speaks directly to the extent to which the GMA-CWB relationship is best understood as dependent on personality.

Finally, we probe all interactions by calculating 95% confidence intervals (CIs) using the Johnson-Neyman (JN) technique (Bauer & Curran, 2005). Results of simple slopes are also shown for illustrative purposes. Alpha values in JN technique and p-values in simple slopes tests were Bonferroni corrected to account for the additional significance tests (Bauer & Curran, 2005).

Results

CWB Predicted by Stability

Tables 3 and 4 show descriptive statistics for all predictor and criteria variables in samples 1 and 2, respectively. Tables 5 and 6 show correlations for all predictor and criteria variables in samples 1 and 2, respectively. In both datasets, CWB was heavily skewed toward low reports of CWB (i.e., skewness of CWB was 0.42 in Sample 1 and 0.46 in Sample 2, indicating a positive, right-skewed distribution in both samples).[4] As noted above, positive skew is typical of CWB data (Guay et al., 2016; Penney & Spector, 2005; van Zyl & de Bruin, 2018). To test our hypothesis that the meta-trait stability moderates the GMA-CWB relationship, we ran moderated multiple regression models for overall CWB predicted from stability and GMA using MLR estimation to account for non-normal distribution. Notably, results suggest that effect sizes in both samples were very close to or exceeded those the study was powered to detect based on sensitivity analyses.

Table 3 Descriptive statistics for variables in Sample 1

Variable	Scoring	*M*	*SD*	Min	Max
GMA	2PL	0.00	0.85	-2.81	1.94
	sum	0.61	0.17	0.00	1.00
Conscientiousness	GGUM	0.00	0.95	-2.98	2.11
	sum	3.74	0.75	1.22	5.00
Agreeableness	GGUM	0.00	0.92	-2.92	2.25
	sum	3.71	0.64	1.60	5.00
Emotional stability	GGUM	0.00	0.94	-3.62	3.47
	sum	3.41	0.92	1.00	5.00
Stability	GGUM composite	0.00	0.75	-2.56	2.45
	sum composite	0.00	0.76	-2.55	1.82
CWB-I	GRM	0.00	0.89	-1.03	3.58
	sum	1.85	1.08	1.00	7.00
CWB-O	GRM	0.00	0.91	-1.60	2.87
	sum	2.11	0.92	1.00	6.00
CWB	GRM composite	0.00	0.90	-1.45	3.48
	sum composite	0.00	0.91	-1.00	4.10

Composite variables were created by z-scoring and then averaging component variables. $N = 1{,}412$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *2PL* Two-parameter logistic model; *GGUM* Generalized graded unfolding model; *GRM* Graded response model; *Sum* Traditional sum-scoring approach (average); *Min.* Minimum; *Max.* Maximum

Results of moderated multiple regression models for overall CWB predicted from stability and GMA are shown in Table 7 for both samples. Figure 2 illustrates these results in both samples. In Sample 1, results showed a significant interactive effect of stability and GMA on overall CWB ($b = 0.07$, $p = 0.002$). In Sample 2, results showed a significant interactive effect of stability and GMA on overall CWB ($b = 0.14$, $p < 0.001$). Thus, both samples showed evidence of a significant GMA-stability interaction for overall CWB. As noted above, researchers have previously argued that although ΔR^2 is traditionally used as an index of interaction effect size, it yields misleadingly small effect sizes that make interpretation difficult. Indeed, the values of ΔR^2 found here (see Table 7) are slightly larger than the median effect size for interactions reported by Aguinis et al. (2005). Thus, we report the recently proposed index $\Delta R^2_{mo.}$ We also report total R^2 for context. In Sample 1, inclusion of the meta-trait stability accounted for 77.8% of the variability in the relationship between GMA and CWB (i.e., ΔR^2_{mo}). The total proportion of variance explained by the model (i.e., R^2) was 27.5%. In Sample 2, stability accounted for 35.2% percent of the variability in the relationship between GMA and CWB,

[4] We also ran a z-test to compare the proportion of "never" endorsement in Sample 1 (7.30%) and Sample 2 (9.36%) and found no significant difference, $z = -1.92$, $p = .059$.

Table 4 Descriptive statistics for variables in Sample 2

Variable	Scoring	*M*	*SD*	Min	Max
GMA (spatial)	2PL	0.00	0.84	-1.98	1.65
	sum	0.51	0.22	0.00	1.00
GMA (verbal)	2PL	0.00	0.76	-1.77	1.53
	sum	0.53	0.23	0.00	1.00
GMA	2PL composite	0.00	0.88	-2.33	1.98
	sum composite	0.00	0.87	-2.33	2.15
Conscientiousness	GGUM	0.00	0.95	-2.22	4.85
	sum	4.49	0.77	2.12	6.00
Agreeableness	GGUM	-0.01	0.94	-2.19	2.68
	sum	4.26	0.69	1.83	6.00
Emotional Stability	GGUM	0.00	0.96	-2.71	3.06
	sum	4.02	0.93	1.42	6.00
Stability	GGUM composite	0.00	0.81	-2.44	2.76
	sum composite	0.00	0.83	-2.76	2.09
CWB-I	GRM	0.00	0.92	-1.03	2.44
	sum	2.06	1.41	1.00	6.71
CWB-O	GRM	0.00	0.95	-1.63	2.48
	sum	2.44	1.30	1.00	6.58
CWB	GRM composite	0.00	0.94	-1.43	2.47
	sum composite	0.00	0.96	-0.93	3.10

Composite variables were created by z-scoring and then averaging component variables. $N = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *2PL* Two-parameter logistic model; *GGUM* Generalized graded unfolding model; *GRM* Graded response model; *Sum* Traditional sum-scoring approach (average); *Min.* Minimum; *Max.* Maximum

and the total proportion of variance explained by the model was 17.3%.

Additionally, because our hypothesis specified that the effect of GMA on overall CWB would be negative among low- but not high-stability results, we calculated the JN CIs to determine at what levels of stability the effect of GMA on CWB was significant (Table 8). The effect of GMA on CWB was significant outside the JN 95% CI [-0.58, 1.54] in Sample 1 and [0.80, 1.87] in Sample 1. Specifically, as illustrated in Fig. 2, the effect of GMA on CWB was negative and significant at low levels of stability across both samples.

CWB Subdimensions Predicted by Stability

RQ1 inquired whether the interactive effect of GMA and stability may differ by CWB subdimension. Results of moderated multiple regression models for CWB subdimensions predicted from stability and GMA are shown in Table 9 for samples 1 and 2. Figure 3 illustrates these results in Sample 1, and Fig. 4 illustrates these results in Sample 2. In Sample 1, the results showed a significant interactive effect of stability and GMA on CWB-I ($b = 0.06$, $p = 0.011$) as well as CWB-O ($b = 0.07$, $p = 0.005$). In Sample 2, results showed a significant interactive effect of stability and GMA on CWB-I ($b = 0.15$, $p < 0.001$) as well as CWB-O ($b = 0.12$, $p < 0.001$).

In Sample 1, inclusion of the meta-trait stability accounted for 45.9% of the variability in the relationship between GMA and CWB-I ($R^2 = 0.114$) and accounted for 99.5% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.171$). In Sample 2, stability accounted for 33.5% percent of the variability in the relationship between GMA and CWB-I ($R^2 = 0.230$) and accounted for 37.6% percent of the variability in the relationship between GMA and CWB-O ($R^{2=}$ 0.256). Table 10 shows JN 95% CIs and results of simple slopes tests for the effect of GMA on CWB subdimensions in both samples. For CWB-I, the effect of GMA on CWB was significant outside the JN 95% CI [-0.12, 7.91] in Sample 1 and [0.82, 1.97] in Sample 2 such that there was a negative effect of GMA on CWB-I at low levels of stability across both samples.

In contrast, results for CWB-O differed across samples. In Sample 1, the effect of GMA on CWB-O was significant outside the JN 95% CI [-2.37, 1.06] such that there was no negative effect of GMA on CWB at low levels of stability.

Table 5 Correlations for variables in Sample 1

Variable	1	2	3	4	5	6	7
1. GMA							
2. Conscientiousness	-0.04						
3. Agreeableness	0.10***	0.37***					
4. Emotional stability	0.00	0.38***	0.26***				
5. Stability	0.03	0.78***	0.73***	0.73***			
6. CWB-I	-0.07**	-0.22***	-0.36***	-0.15***	-0.33***		
7. CWB-O	-0.02	-0.39***	-0.29***	-0.23***	-0.41***	0.62***	
8. CWB (overall)	0.05	-0.34***	-0.36***	-0.21***	-0.41***	0.90***	0.90***

$N = 1{,}412$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB

$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$

Table 6 Correlations for variables in Sample 2

Variable	1	2	3	4	5	6	7	8	9
1. GMA (spatial)									
2. GMA (verbal)	0.54***								
3. GMA	0.88***	0.88***							
4. Consc	0.17***	0.07**	0.14***						
5. Agr	0.08**	0.03	0.06*	0.46***					
6. ES	0.07**	0.06*	0.08**	0.62***	0.36***				
7. Stab	0.13***	0.07*	0.11***	0.86***	0.75***	0.82***			
8. CWB-I	-0.28***	-0.15***	-0.24***	-0.45***	-0.27***	-0.28***	-0.41***		
9. CWB-O	-0.21***	-0.13***	-0.20***	-0.53***	-0.25***	-0.36***	-0.47***	0.76***	
10. CWB	-0.26***	-0.15***	-0.23***	-0.52***	-0.28***	-0.34***	-0.47***	0.94***	0.94***

$N = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *Stab*. Stability; *Consc*. Conscientiousness; *Agr*. Agreeableness; *ES* Emotional stability

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$

Table 7 Results of regression analyses for CWB predicted by stability and GMA

Variable	Sample 1			Sample 2		
	CWB			CWB		
	b	SE	*p*	*b*	SE	*p*
Intercept	0.00	0.02	0.470	-0.02	0.02	0.254
GMA	-0.03	0.03	0.161	-0.17	0.02	< 0.001
Stab	-0.40	0.02	< 0.001	-0.44	0.02	< 0.001
Stab. x GMA	0.07	0.02	0.002	0.14	0.03	< 0.001
R^2 (ΔR^2)	0.173 (0.005)			0.275 (0.018)		
ΔR^2_{mo}	0.778			0.352		

Models were also run in which quadratic terms were included for all predictors per Cortina (1993). In no case did inclusion of the quadratic term change the significance of the interaction term. Consistent with directional hypothesis testing, p-values are one-tailed. $N_1 = 1{,}412$; $N_2 = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *Stab*. Stability

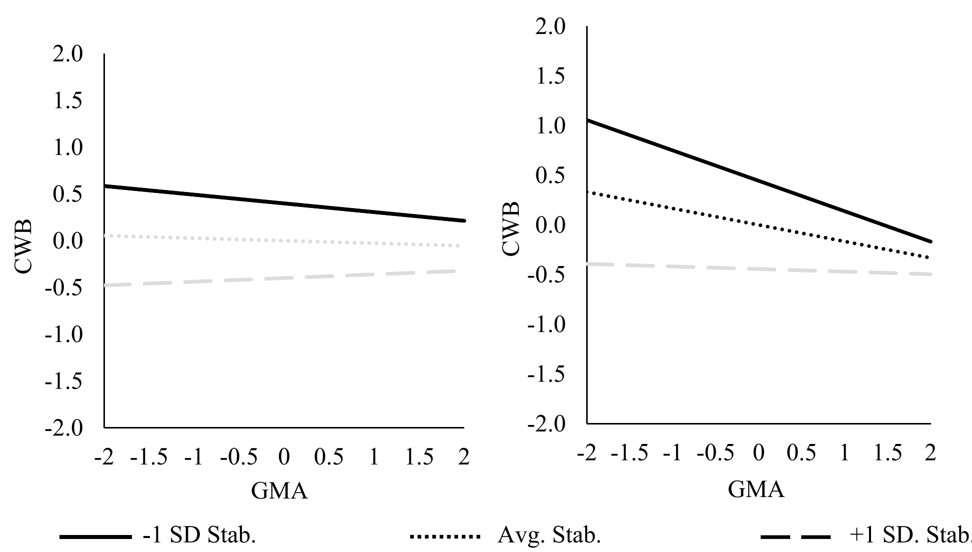

Fig. 2 Effect of GMA on CWB moderated by stability in sample 1 (left) and sample 2 (right). *Note*. Johnson-Neyman 95% CI [-0.58, 1.54] in Sample 1 and [0.80, 1.87] in Sample 2; effect of GMA on CWB is signifciant at stability levels outside the confidence interval. Black lines indicate significant simple slopes. Gray lines indicate non-significant simple slopes. Stab. = stability. GMA = general mental ability. CWB = counterproductive work behavior

Table 8 Results of simple slope tests for the effect of GMA on CWB at low, average, and high levels of stability

Level of Stability	Sample 1			Sample 2		
	CWB			CWB		
	b	SE	*p*	*b*	SE	*p*
Low (-1 SD)	-0.09	0.04	0.019	-0.31	0.04	< 0.001
Average	-0.03	0.03	0.161	-0.17	0.02	< 0.001
High (+1 SD)	0.04	0.03	0.201	-0.03	0.03	0.397
95% CI	[-0.58, 1.54]			[0.80, 1.87]		

Note. 95% confidence interval applies JN technique to continuous variables (see Bauer & Curran, 2005). When the respective personality trait is outside the confidence interval, the slope of general mental ability (GMA) is significant. Low indicates 1 standard deviation below the mean; high indicates one standard deviation above the mean. JN alpha and simple slopes p-values are Bonferroni corrected (Bauer & Curran, 2005). $N_1 = 1{,}412$; $N_2 = 1{,}241$; *CWB* Counterproductive work behavior

Table 9 Results of regression analyses for CWB-I and CWB-O predicted by stability and GMA

Variable	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Sample 1						
Intercept	0.00	0.03	0.952	0.00	0.02	0.941
GMA	-0.06	0.03	0.041	0.01	0.03	0.840
Stab	-0.32	0.02	< 0.001	-0.40	0.03	< 0.001
Stab. x GMA	0.06	0.02	0.011	0.07	0.02	0.005
R^2 (ΔR^2)	0.114 (0.003)			0.171 (0.005)		
ΔR^2_{mo}	0.459			0.995		
Sample 2						
Intercept	-0.02	0.03	0.501	-0.01	0.02	0.590
GMA	-0.18	0.02	< 0.001	-0.13	0.03	< 0.001
Stab	-0.38	0.02	< 0.001	-0.45	0.02	< 0.001
Stab. x GMA	0.15	0.02	< 0.001	0.12	0.03	< 0.001
R^2 (ΔR^2)	0.230 (0.020)			0.256 (0.012)		
ΔR^2_{mo}	0.335			0.376		

Models were also run in which quadratic terms were included for all predictors per Cortina (1993). In no case did inclusion of the quadratic term change the significance of the interaction term. P-values are two-tailed. $N_1 = 1{,}412$; $N_2 = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *Stab*. Stability

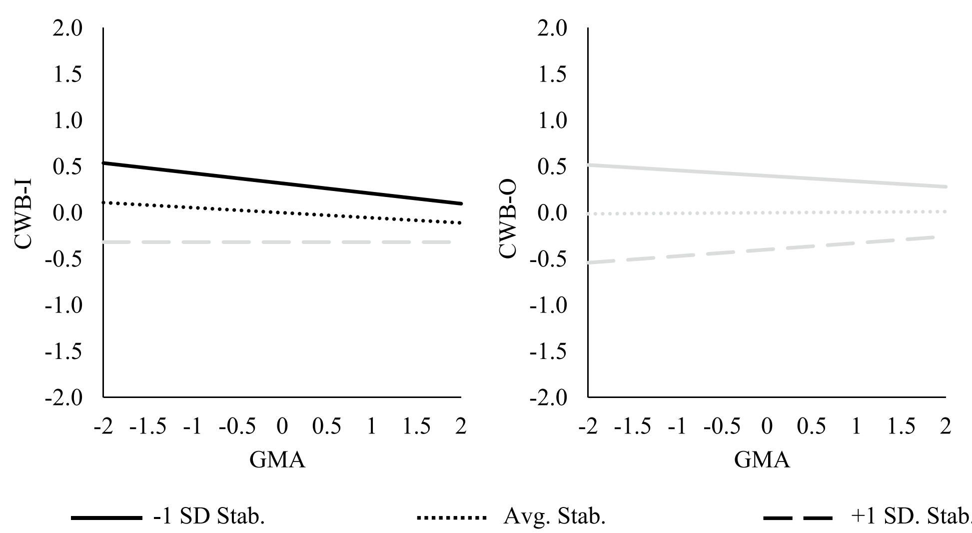

Fig. 3 Effect of GMA on CWB-I (left) and CWB-O (right) moderated by stability in sample 1**.** *Note*. Johnson-Neyman 95% CI [-0.12, 7.91] in Sample 1 and [-2.37, 1.06] in Sample 2; effect of GMA on CWB is signifciant at stability levels outside the confidence interval. Black lines indicate significant simple slopes. Gray lines indicate non-significant simple slopes. Stab. = stability. GMA = general mental ability. CWB = counterproductive work behavior. CWB-I = interpersonal CWB. CWB-O = organizational CWB

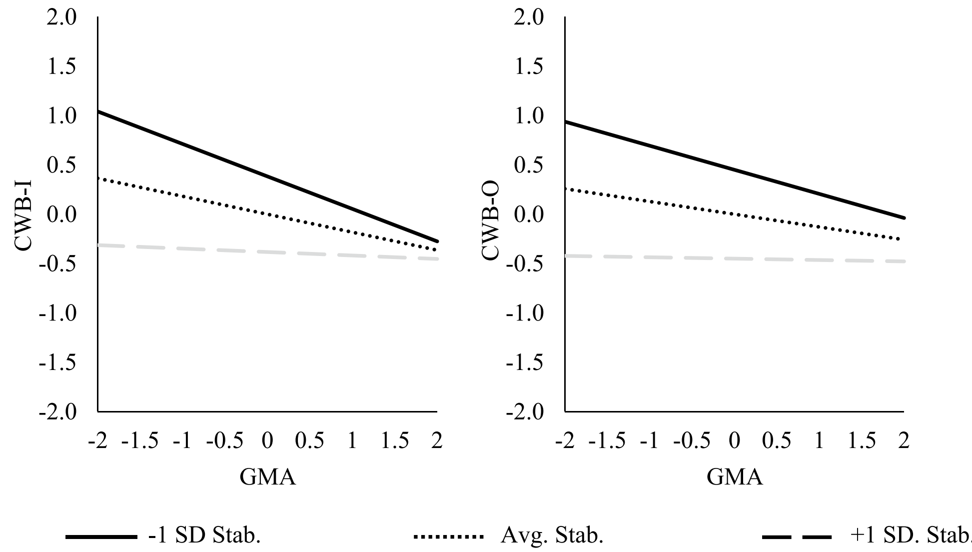

Fig. 4 Effect of GMA on CWB-I (left) and CWB-O (right) moderated by stability in sample 2**.** *Note*. Johnson-Neyman 95% CI [0.82, 1.97] in Sample 1 and [0.57, 2.55] in Sample 2; effect of GMA on CWB is signifciant at stability levels outside the confidence interval. Black lines indicate significant simple slope. Gray lines indicate non-significant simple slopes. Stab. = stability. GMA = general mental ability. CWB = counterproductive work behavior. CWB-I = interpersonal CWB. CWB-O = organizational CWB

Table 10 Results of simple slope tests for the effect of GMA on CWB-I and CWB-O at low, average, and high levels of stability

	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Sample 1						
Low (-1 SD)	-0.11	0.04	0.006	-0.06	0.04	0.264
Average	-0.06	0.03	0.042	0.01	0.03	0.840
High (+1 SD)	0.00	0.03	1.000	0.07	0.03	0.058
95% CI	[-0.12, 7.91]			[-2.37, 1.06]		
Sample 2						
Low (-1 SD)	-0.33	0.04	< 0.001	-0.24	0.04	< 0.001
Average	-0.18	0.02	< 0.001	-0.13	0.03	< 0.001
High (+1 SD)	-0.04	0.03	0.482	-0.01	0.04	1.000
95% CI	[0.82, 1.97]			[0.57, 2.55]		

95% confidence interval applies JN technique to continuous variables (see Bauer & Curran, 2005). When the respective personality trait is outside the confidence interval, the slope of general mental ability (GMA) is significant. Low indicates 1 standard deviation below the mean; high indicates one standard deviation above the mean. JN alpha and simple slopes p-values are Bonferroni corrected (Bauer & Curran, 2005). *P*-values are two-tailed. $N_1 = 1{,}412$; $N_2 = 1{,}241$; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB. *CWB-O* Organizational CWB

Rather, there was a *positive* effect of GMA on CWB at high levels of stability. In Sample 2, the effect of GMA on CWB-O was significant outside the JN 95% CI [0.57, 2.55] such that there was a negative effect of GMA on CWB at low levels of stability. Thus, whereas results for CWB-I were generally consistent across the two samples and consistent with the form of the interaction hypothesized for overall CWB, results for CWB-O differed by sample.

CWB Predicted by Lower-order Traits

To explore RQ2 regarding whether traits comprising the meta-trait stability (conscientiousness, agreeableness, and emotional stability) differentially moderated the GMA-CWB relationship, moderated multiple regression analyses were also conducted for CWB at the trait-level in both samples. Table 11 shows results of these moderated multiple regression models for CWB predicted from GMA and conscientiousness, agreeableness, and emotional stability in samples 1 and 2. Figures 5 and 6 illustrate these interactions in Sample 1 and Sample 2, respectively. Table 12 shows JN 95% CIs and results of simple slopes tests for the effect of GMA on CWB in both samples. Trait-level results for CWB subdimensions are included in Appendix 1.

Conscientiousness In Sample 1, there was a significant interactive influence of conscientiousness and GMA on overall CWB ($b = 0.06$, $p = 0.021$), and conscientiousness

Table 11 Results of regression analyses for CWB predicted by lower-order traits and GMA

Variable	Sample 1			Sample 2		
	CWB			CWB		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Intercept	0.00	0.03	0.930	-0.02	0.02	0.433
GMA	-0.05	0.03	0.067	-0.15	0.02	< 0.001
Consc	-0.34	0.03	< 0.001	-0.49	0.02	< 0.001
Consc. x GMA	0.06	0.03	0.021	0.14	0.02	< 0.001
R^2 (ΔR^2)	0.123 (0.004)			0.318 (0.016)		
ΔR^2_{mo}	0.502			0.379		
Agreeableness						
Intercept	-0.01	0.03	0.814	-0.01	0.03	0.771
GMA	0.00	0.03	0.880	-0.21	0.03	< 0.001
Agr	-0.36	0.03	< 0.001	-0.29	0.03	< 0.001
Agr. x GMA	0.06	0.03	0.021	0.12	0.03	< 0.001
R^2 (ΔR^2)	0.136 (0.004)			0.137 (0.013)		
ΔR^2_{mo}	0.955			0.209		
Emotional Stability						
Intercept	0.00	0.03	0.998	-0.01	0.03	0.663
GMA	-0.04	0.03	0.155	-0.19	0.03	< 0.001
ES	-0.20	0.03	< 0.001	-0.33	0.03	< 0.001
ES x GMA	0.07	0.03	0.008	0.15	0.03	< 0.001
R^2 (ΔR^2)	0.053 (0.005)			0.181 (0.020)		
ΔR^2_{mo}	0.692			0.317		

Models were also run in which quadratic terms were included for all predictors per Cortina (1993). In no case did inclusion of the quadratic term change the significance of the interaction term. P-values are two-tailed. $N_1 = 1{,}412$; $N_2 = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *Consc*. Conscientiousness; *Agr*. Agreeableness; *ES* Emotional stability

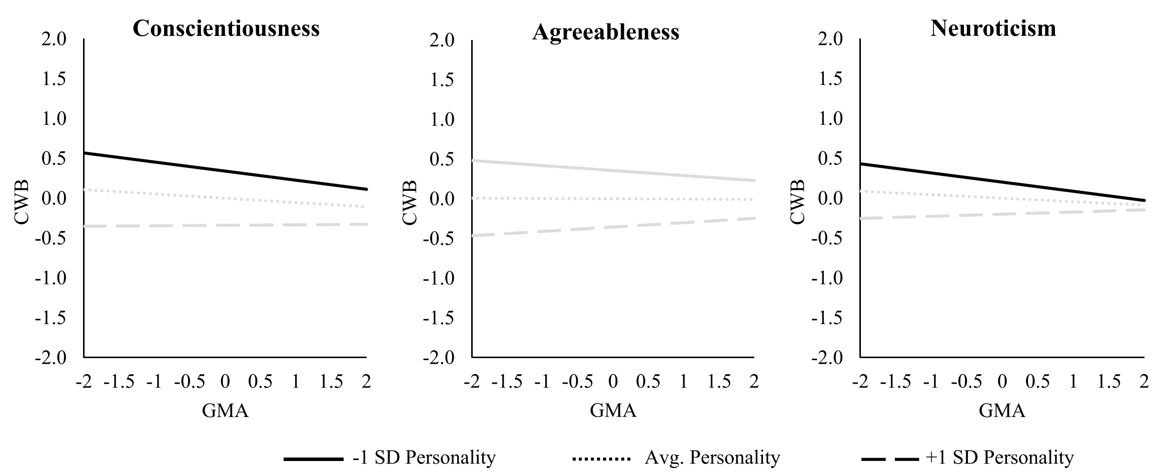

Fig. 5 Effect of GMA on CWB moderated by Conscientiousness (left), Agreeableness (middle), and Emotional Stability (right) in Sample 1**.** *Note*. Johnson-Neyman 95% CI [-0.28, 24.03] for conscientiousness, [-10.08, 2.05] for agreeableness, [-0.50, 4.01] for emotional stability; effect of GMA on CWB is signifciant at stability levels outside the confidence interval. Black lines indicate significant simple slope. Black lines indicate significant simple slope. Gray lines indicate non-significant simple slopes. GMA = general mental ability. CWB = counterproductive work behavior

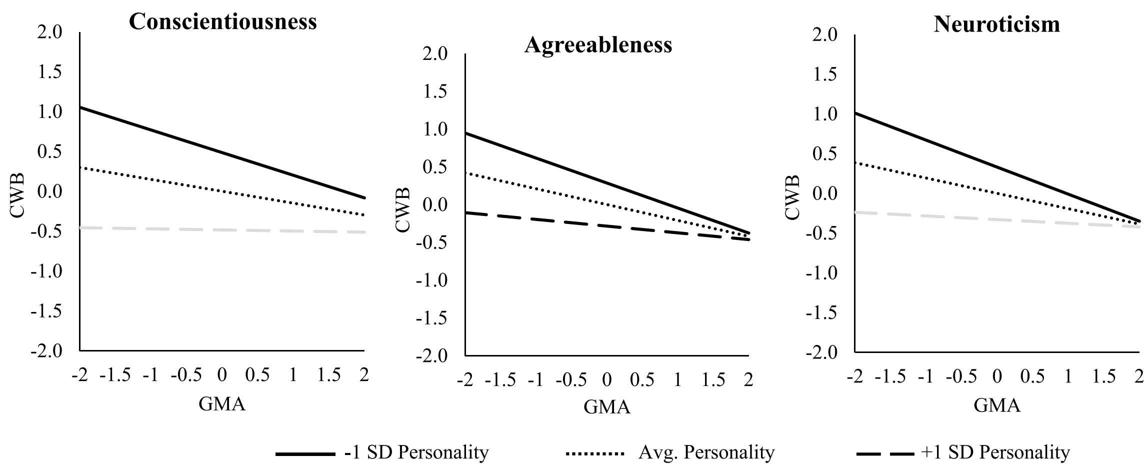

Fig. 6 Effect of GMA on CWB Moderated by Conscientiousness (left), Agreeableness (middle), and Emotional Stability (right) in Sample 2**.** *Note*. Johnson-Neyman 95% CI [0.67, 1.89] for conscientiousness, [1.04, 3.08] for agreeableness, [0.83, 2.32] for emotional stability; effect of GMA on CWB is signifciant at stability levels outside the confidence interval. Black lines indicate significant simple slope. Gray lines indicate non-significant simple slopes. GMA = general mental ability. CWB = counterproductive work behavior

Table 12 Results of simple slope tests for GMA on CWB at low, average, and high levels of lower-order traits

	Sample 1			Sample 2		
	CWB			CWB		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Low (-1 SD)	-0.12	0.04	0.014	-0.28	0.04	< 0.001
Average	-0.05	0.03	0.067	-0.15	0.02	< 0.001
High (+ 1 SD)	0.01	0.04	1.000	-0.02	0.03	1.000
95% CI	[-0.28, 24.03]			[0.67, 1.89]		
Agreeableness						
Low (-1 SD)	-0.06	0.04	0.286	-0.33	0.05	< 0.001
Average	0.00	0.03	0.880	-0.21	0.03	< 0.001
High (+ 1 SD)	0.06	0.03	0.180	-0.09	0.04	0.030
95% CI	[-10.08, 2.05]			[1.04, 3.80]		
Emotional stability						
Low (-1 SD)	-0.12	0.04	0.018	-0.34	0.04	< 0.001
Average	-0.04	0.03	0.156	-0.19	0.03	< 0.001
High (+ 1 SD)	0.03	0.04	0.936	-0.05	0.04	0.362
95% CI	[-0.50, 4.01]			[0.83, 2.32]		

95% confidence interval applies JN technique to continuous variables (see Bauer & Curran, 2005). When the respective personality trait is outside the confidence interval, the slope of general mental ability (GMA) is significant. Low indicates 1 standard deviation below the mean; high indicates one standard deviation above the mean. JN alpha and simple slopes p-values are Bonferroni corrected (Bauer & Curran, 2005). P-values are two-tailed. $N_1 = 1{,}412$; $N_2 = 1{,}241$; *CWB* Counterproductive work behavior

accounted for 50.2% of the variability in the relationship between GMA and CWB ($R^2 = 0.123$). In Sample 2, conscientiousness and GMA showed a significant interactive effect on CWB ($b = 0.14$, $p < 0.001$), and conscientiousness accounted for 37.9% of the variability in the relationship between GMA and CWB ($R^2 = 0.318$). The effect of GMA on CWB was significant outside the JN 95% CI [-0.28, 24.03] in Sample 1 and [0.67, 1.89] in Sample 2 such that across both samples, the effect of GMA on CWB was negative at low levels of conscientiousness. Thus, across both samples, the form of the GMA-conscientiousness interaction was generally consistent with the form of the hypothesized GMA-stability interaction.

Agreeableness In Sample 1, there was a significant interactive influence of agreeableness and GMA on overall CWB ($b = 0.06$, $p = 0.021$), and agreeableness accounted for 95.5% of the variability in the relationship between GMA and CWB ($R^2 = 0.136$). In Sample 2, agreeableness and GMA showed a significant interactive effect on CWB ($b = 0.12$, $p < 0.001$). Agreeableness accounted for 20.9% of the variability in the relationship between GMA and CWB ($R^2 = 0.137$). The effect of GMA on CWB was significant outside the JN 95% CI [-10.08, 2.05] in Sample 1 such that the effect of GMA on CWB was *positive* at very high levels of agreeableness. In contrast, the effect of GMA on CWB was significant outside the JN 95% CI [1.04, 3.80] in Sample 2, indicating that the effect of GMA on CWB was *negative* at low levels of agreeableness. Thus, the form of the GMA-agreeableness interaction was consistent with the form of the hypothesized GMA-stability interaction in Sample 2 but not in Sample 1.

Emotional Stability In Sample 1, there was a significant interactive influence of emotional stability on overall CWB ($b = 0.07$, $p = 0.008$), and emotional stability accounted for 69.2% of the variability in the relationship between GMA and CWB ($R^2 = 0.053$). In Sample 2, emotional stability and GMA also showed a significant interactive effect on CWB ($b = 0.15$, $p < 0.001$), and emotional stability accounted for 31.7% of the variability in the relationship between GMA and CWB ($R^2 = 0.181$). The effect of GMA on CWB was significant outside the JN 95% CI [-0.50, 4.01] in Sample 1 and [0.83, 2.32] in Sample 2 such that the effect of GMA on CWB was negative at low levels of emotional stability. Thus, across both samples, the form of the interaction between GMA and emotional stability was generally consistent with the form of the hypothesized GMA-stability interaction.

Discussion

Although researchers have commonly theorized that GMA *should* negatively predict CWB, empirical research has generally not supported such a relationship. Lack of empirical support for a negative GMA-CWB relationship is particularly surprising given widespread support for a negative link between GMA and deviance generally. In the current study, we aimed to clarify the relationship between GMA and CWB by reframing the inhibitory effect of GMA on CWB as conditional on the meta-trait of stability.

Overall, results across both samples support the hypothesis. Results showed a significant interactive effect of stability and GMA on overall CWB in both samples such that GMA showed a negative relationship with CWB among individuals with low but not high stability. That is, GMA only inhibits CWB among individuals who have dispositional tendencies to engage in CWB. Further, we utilized a new index of moderation effect size, ΔR^2_{mo}, that suggests a substantial portion of the variance in the GMA-CWB relationship is accounted for by these moderating effects. Thus, results further suggest that prior research may not have found consistent support for a negative GMA-CWB relationship because it did not account for an interaction with key personality traits.

In addition to hypothesizing about the interactive influence of stability and GMA on overall CWB, we also posed two research questions. In RQ1 we investigated whether results may differ by CWB subdimension. Across both samples, results showed a significant interactive effect of stability and GMA on CWB-I that was consistent with the hypothesized interaction. However, results for CWB-O differed across the two samples. In Sample 1, results did not show a negative relationship between GMA and CWB-O among low stability individuals as hypothesized. Rather, the JN 95% CI revealed that at stability levels higher than 1.06 SD above mean stability, the effect of GMA on CWB-O was positive. However, these results were not replicated in Sample 2. In Sample 2, there was a significant interactive effect of stability and GMA on CWB-O consistent with the hypothesized interaction.

Further, in answer to RQ2, we explored the potential for differential interactions by lower-order traits (i.e., conscientiousness, agreeableness, and emotional stability). Regarding overall CWB, conscientiousness and emotional stability showed significant interactions with GMA in both samples consistent with the hypothesized form. In contrast, although there was evidence of a significant interactive influence of agreeableness and GMA on overall CWB in both samples, only Sample 2 was consistent with the form hypothesized. The JN 95% CI revealed that the effect of GMA on CWB in Sample 1 was only significant at very high levels of agreeableness, above 2.05 SD above the mean of agreeableness.

One reviewer expressed interest in further understanding differences in the CWB subdimensions, including whether these results may be clarified at the trait level. In Appendix 1, we provide results of moderated regression analyses for both CWB subdimensions predicted by an interaction between GMA and all three lower-order traits. For CWB-I, Sample 1 trait-level results demonstrated significant GMA interactions with both conscientiousness and emotional stability consistent with the form hypothesized but no significant interaction with agreeableness. In Sample 2, there was a significant interaction between GMA and all three lower-order traits for CWB-I.

In contrast, trait-level results for CWB-O differed by sample. In Sample 1, there was no significant conscientiousness-GMA interaction for CWB-O. Sample 1 results supported interactions for agreeableness and emotional stability but not of the form hypothesized. Rather, results suggested a possible positive influence of GMA on CWB-O at high levels of agreeableness and emotional stability. Because there was no agreeableness-GMA interaction for CWB-I, and overall CWB is an average of CWB-I and CWB-O, these results explain why there was some evidence of a positive influence of GMA on overall CWB for high agreeableness. However, these findings were not replicated in Sample 2. All trait-level CWB-O analyses in Sample 2 supported the hypotheses. Further, interaction effect sizes for CWB-O were relatively consistent across the lower-order traits in Sample 2, suggesting that the meta-trait interaction cannot be attributed to any one lower-order trait. Thus, although results for CWB-O were somewhat inconsistent *between samples*, results did not notably differ *between traits*. That is, none of the trait-level interactions were consistent with the hypothesized form for CWB-O in Sample 1, but all of the trait-level interactions were consistent with the hypothesized form for CWB-O in Sample 2. Consequently, we believe inconsistencies for CWB-O in Sample 1 should be interpreted with caution.

Ultimately, we believe that trait-level results emphasize the relevance of the meta-trait *stability* for understanding the inhibitory effect of GMA. Although prior research has empirically supported negative relationships between all three lower-order traits and CWB, the current study is the first to test a relationship between the meta-trait stability and CWB. Thus, we proposed and investigated RQ2 to allow for the possibility that a stability-GMA interaction effect on CWB may be primarily driven by only one or two of the three lower-order traits. However, for overall CWB, 5 of the 6 trait-level interactions tested were of the same form as the hypothesized stability-GMA interaction. Thus, we believe our results are consistent with Marcus and Schuler (2004)'s findings, which suggest that a broad form of self-control–such as the meta-trait stability used here–is more appropriate than narrower dimensions when predicting general CWB.

A reviewer also noted that there were some differences between the two samples that may account for the generally stronger effects seen in Sample 2 relative to Sample 1. First, mean sum-scores for overall CWB and both subtypes were higher in Sample 2 relative to Sample 1. These mean differences may stem in part from a higher proportion of males in Sample 2. Moreover, Sample 1 was collected in May 2016 whereas Sample 2 was collected in March 2020, just after the beginning of the COVID-19 pandemic. Finally, the two samples showed slightly different patterns of GMA-personality correlations. However, with the possible exception of conscientiousness in Sample 2, correlations were generally within the range reported in prior meta-analyses of the GMA-personality relationship (Ackerman & Heggestad, 1997).

Despite these differences between Sample 1 and Sample 2, results were generally consistent between the two samples. Although there were some differences between the two samples by CWB subdimensions and at the trait-level, result for our hypothesis—which concerned the interactive influence of GMA and the meta-trait stability on overall CWB—were consistent across both samples. In fact, rather than call into question the generalizability of findings, we believe that differences across samples underscore the robustness of our findings.

Implications for Theory and Practice

Results of the current study have important implications for both theory and practice. Despite ample support for a negative GMA-deviance relationship in the criminology literature, prior research has surprisingly not supported a GMA-CWB relationship. The current study clarifies the theoretical relationship between GMA and CWB by revisiting the rationale behind the popular inhibitory effect. Prior empirical studies have operationalized an inhibitory effect as a main effect of GMA on CWB. Here, we argue for a conditional inhibitory effect such that GMA shows an inhibitory effect on CWB among those with impulsive tendencies (i.e., low to moderate levels of the meta-trait stability). In contrast, GMA does not show a negative relationship with CWB among those who do not have such impulsive tendencies. Thus, our results are consistent with a conditional inhibitory effect such that GMA helps individuals with dispositional tendencies to engage in CWB to anticipate the consequences of their actions and inhibit those impulses.

Researchers have also proposed explanations for why prior empirical findings have not yet supported a negative GMA relationship with CWB, and our results speak to some of these explanations. As mentioned above, one explanation is that there are differential effects of GMA by subdimensions. Overall, our results suggest inconsistent evidence for differential effects by CWB subdimensions. Sample 1 showed evidence of possible differential interaction effects for the CWB subdimensions such that CWB-I results were consistent with the form of the interaction hypothesized for overall CWB, but CWB-O results were not. However, there was no support for differential effects by CWB subdimension in Sample 2, so differences in Sample 1 should be interpreted with caution.

Another commonly proposed explanation for the lack of support for a GMA-CWB relationship is differential detection. The use of self-reports was an intentional study design choice to prevent confounding differential engagement in CWB with differential detection of CWB. Thus, the current results suggest that GMA has an inhibitory effect on actual *engagement* in CWB. Conversely, results suggest a negative GMA-CWB relationship cannot be fully attributed to lower detection of CWBs committed by high-GMA individuals.

Results also have practical implications for the use of GMA to predict CWB. The current results were consistent with prior research that suggests personality traits show stronger relationships with CWB than does GMA (Gonzalez-Mulé et al., 2014). However, results also suggest that incorporating a stability-GMA interaction into selection models may be helpful, particularly when attempting to screen out applicants (i.e., the applicants most likely to engage in CWB). To demonstrate the potential benefit of including a stability-GMA interaction in a selection system, we conducted a simulated selection analysis in which we compared a stability-GMA interaction model and a stability-only model for screening out the bottom 10% and 20% of applicants see Appendix 2 for details). Further, we cross-validated parameters by using coefficients estimated from each sample to predict CWB in both samples separately, as well as the combined samples. In nearly all cases (22 of the 24 comparisons summarized in Appendix 2), the stability-interaction model outperformed the stability-only model. That is, the interaction model predicted or "screened out" the highest CWB applicants with greater accuracy than did the stability-only model. If selection practitioners are interested in screening out applicants most likely to engage in CWB, the interaction model is likely to offer enhanced utility.

Results also have implications for selection practitioners interested in limiting GMA-related adverse impact. Prior research suggests that personality tests have reduced adverse impact relative to GMA measures (Foldes et al., 2008; Hough et al., 2001). Our results suggest that if applicants are first selected for high stability, then there is no additional predictive utility in selecting on GMA at later stages, because the GMA-CWB relation is null for those high in stability). However, if the range of stability in the applicant pool is restricted such that the applicant pool includes very few high-stability individuals, then GMA may help to decrease the occurrence of future CWB. Notably, posthoc analyses of our data suggested race differences in predicted CWB were smaller when utilizing stability in addition to GMA as opposed to using GMA alone, and inclusion of the interaction term had no desirable or undesirable influence on these differences.[5] Consistent with best practices in selection systems generally, practitioners should calculate utility estimates that consider the added value of GMA assessments.

Limitations and Future Directions

In the current paper, we focused on reconceptualizing the inhibitory effect as an *interactive* influence of GMA on CWB and showed support for a GMA-stability interaction. As mentioned above, a variety of other explanations have been proposed for why prior empirical studies have not shown clean support for a negative GMA-CWB relationship. Importantly, an inhibitory interactive effect is not mutually exclusive of these other explanations, and we encourage future researchers to continue to explore these possibilities.

For example, our results suggest that high GMA leads to differential *engagement in*, not detection of, CWBs. However, because we utilized only self-reported CWBs, we also cannot rule out differential detection in objective and other-reported CWBs. That is, it is possible that highly intelligent individuals have the double advantage of being both

[5] In our first dataset (78% White, 6% Black, 5% Asian, 5% Hispanic) we found no significant observed differences in CWB between race categories, $F(3,1314) = 0.56$, $p = .640$. However, predicting CWB from GMA scores alone resulted in a significant race effect, $F(3,1314) = 4.72$, $p = .002$, such that Black respondents had higher predicted CWB, $t(1176) = 3.77$, $p < .001$, with a moderately large effect size, $d = .44$. Notably, in this sample GMA alone did not significantly predict CWB (see Table 5). However, all other models utilizing stability (stability alone, joint "main effects" of GMA and stability, or interacting GMA and stability) showed no significant differences in predicted CWB between race categories. In our second dataset (74% White, 12% Black, 5% Asian, 5% Hispanic), we found significant observed differences in CWB, $F(3,1177) = 6.99$, $p < .001$, such that Asian respondents showed lower CWB than White respondents, $t(968) = -2.55$, $p = .011$, $d = -.33$, whereas Black respondents showed higher CWB than White respondents, $t(1056) = 3.40$, $p < .001$, $d = .30$. Using GMA alone to predict CWB resulted in a significant effect of race on predicted values of CWB, $F(3,1177) = 21.84$, $p < .001$. However, when using predicted CWB (via GMA), the Black-White effect size was exaggerated at $d = .65$ for predicted CWB, compared to $d = .30$ for observed CWB (see above). The Asian-White effect for predicted CWB was similar to the observed CWB finding at $d = -.36$. Utilizing stability as a lone predictor showed no significant race differences in predicted CWB. However, the effect of race category returned when considering joint prediction models. Using the "main effects" of GMA and stability resulted in a significant race effect, $F(3,1177) = 5.49$, $p < .001$, with significant and similar adverse effects for Black respondents, $d = .31$, and Hispanic respondents, $d = .29$. Nearly identical results were found when utilizing the interaction between GMA and Stability with effect sizes of .29 for Black respondents and .30 for Hispanic respondents.

genuinely less likely to engage in CWBs and less likely to be caught when they do engage in CWBs. Moreover, regarding self-reported *detection* of CWBs, individuals may not be the most accurate informants. Future research should be conducted that considers the influence of a GMA-stability interaction in self- vs. other-reported CWBs.

Another explanation that has been proposed for lack of a GMA-CWB relationship in prior empirical studies is sample heterogeneity (Dilchert et al., 2007). The sample heterogeneity explanation suggests GMA and CWB are related but only in certain occupational groups, and samples that combine occupational groups are likely to obscure this relationship at the population level. However, the rationales for why, when, and how the GMA-CWB relationship should vary by occupation have been generally inconsistent, and there is relatively little empirical support for the sample heterogeneity argument (Gonzalez-Mulé et al., 2014). In the current study, we used a sample with heterogeneous occupational contexts. Thus, our results suggest that heterogeneity is *not* a sufficient explanation for null GMA-CWB relationships in prior studies. Still, our results cannot rule out that the GMA-CWB relationship may differ slightly by sample type. The current study design may have obscured occupational differences in the GMA-CWB relationship, and future researchers might consider possible differential interaction by occupational context when theoretically appropriate.

Similarly, as noted above, the inhibitory effect does not preclude the moral reasoning rationale. In fact, we believe that the inhibitory effect *encompasses* moral reasoning. It is possible that the relationship between GMA and CWB is best represented as moderated mediation in which GMA affects moral reasoning, and moral reasoning in turn interacts with stability to affect CWB. To date, no research of which we are aware has directly tested the effect of moral reasoning on CWB. More research is needed both on the moral reasoning rationale generally and how it may work together with the inhibitory interactive effect identified here.

Future research should also further explore the possibility of a positive GMA-CWB relationship among high stability individuals. The current study shows some evidence for a possible positive effect of GMA on CWB-O at high stability. These results are somewhat surprising and inconsistent with most prior theorizing regarding the relationship between GMA and CWB. One possible explanation for the positive GMA-CWB relationship among high stability individuals is the stressor-emotion model of contextual performance proposed by Spector and colleagues (Spector & Fox, 2002). Although CWB is typically thought of as a maladaptive outcome, this model suggests that CWB may at times function as an adaptive strategy for responding to emotional stressors. Indeed, several studies have found that CWB may be a reaction to stressors (e.g., Bowling & Eschleman, 2010; Krischer et al., 2010; Fox et al., 2001). It is possible that among high stability individuals (i.e., people who are less likely to engage in CWB overall), GMA may be related to one's ability to recognize and adopt effective stress-reducing behaviors. For example, wasting time or intentionally taking long breaks may be interpreted as useful strategies for reducing strain. Still, these positive GMA-CWB relationships and the explanation offered here should be interpreted with caution, particularly given that Sample 2 did not replicate these findings. To directly explore this explanation, future research might directly incorporate the role of workplace stressors into the study design.

Finally, future research should consider the potential influence of a GMA-CWB interaction in deviance generally. The current results suggest that GMA shows a negative relationship with CWB among those with low stability. In contrast, although researchers have found that self-control is an important predictor of deviance (Gottfredson & Hirschi, 1990), empirical findings have largely shown support for a main effect of GMA on deviant behavior irrespective of self-control. Because our reinterpretation of the inhibitory effect is based in the explanation of the effect itself, we would expect a conditional inhibitory effect to also apply to deviance generally. Differences in findings for CWB and deviance generally may be in part attributable to mean differences in stability between the corresponding traditionally studied populations. Future research should test a possible GMA-stability interaction in deviance broadly to further clarify the theoretical relationship of GMA to deviant behavior.

Conclusion

By reinterpreting the commonly theorized inhibitory effect of GMA on CWB as conditional, we have provided one explanation as to why empirical evidence has not generally supported a negative GMA-CWB relationship to date. Results of the current study suggest that GMA only demonstrates a negative relationship among those with dispositional tendencies to engage in CWB. Specifically, GMA inhibits CWB among those with low but not high levels of the meta-trait stability. Future research should consider whether this interactive effect also influences the detection of CWB, whether the form of the interaction varies by occupation, as well as the possibility of mediation via moral reasoning. Finally, future research should explore the implications of a GMA-stability interaction in deviance generally to further delineate the relationship between CWB and broader types of deviance.

Appendix 1

CWB Subdimensions Predicted by Lower-order Traits

Below, we provide results for both CWB-I and CWB-O predicted by conscientiousness, agreeableness, and emotional stability. Table 13 shows results of moderated multiple regression models in Sample 1, and Table 14 shows corresponding results of simple slopes tests. Table 15 shows results of moderated multiple regression models in Sample 2, and Table 16 shows corresponding results of simple slopes tests. Consistent with other exploratory analyses described in the main manuscript, all p-values reported here are two-tailed. Further, JN alpha and p-values for simple slopes tests are Bonferroni corrected (Bauer & Curran, 2005).

Conscientiousness

In Sample 1, there was a significant interactive influence of conscientiousness and GMA on CWB-I ($b = 0.07$, $p = 0.014$) but not CWB-O. Conscientiousness accounted for 41.2% of the variability in the relationship between GMA and CWB-I ($R^2 = 0.058$). The effect of GMA on CWB-I was significant outside the JN 95% CI [0.12, 13.34] such that the effect was negative at low levels of conscientiousness.

In Sample 2, conscientiousness and GMA showed a significant interactive effect on CWB-I ($b = 0.15$, $p < 0.001$) and CWB-O ($b = 0.10$, $p < 0.001$). Conscientiousness accounted for 38.0% of the variability in the relationship between GMA and CWB-I ($R^2 = 0.258$) and 37.7% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.307$). The effect of GMA on CWB-I was significant outside the JN 95% CI [0.71, 1.74] and the effect of GMA on CWB-I was significant outside the JN 95% CI [0.52, 2.68]. Thus, the effect of GMA on both CWB-I and CWB-O was significant and negative at low levels of conscientiousness in Sample 2.

Agreeableness

In Sample 1, there was a significant interactive influence of agreeableness and GMA on CWB-O ($b = 0.07$, $p = 0.008$) but not on CWB-I. Agreeableness accounted for 97.4% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.134$). The effect of GMA on CWB-O was significant outside the JN 95% CI [-4.17, 0.59] such that the effect was *positive* at high levels of agreeableness.

In Sample 2, agreeableness and GMA showed a significant interactive effect on CWB-I ($b = 0.11$, $p < 0.001$) and CWB-O ($b = 0.11$, $p < 0.001$). Agreeableness accounted for 17.1% of the variability in the relationship between GMA and CWB-I ($R^2 = 0.135$) and 25.7% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.108$). The effect of GMA on CWB-I was significant outside the JN 95% CI [1.21, 4.26] and the effect of GMA on CWB-O was significant outside the JN 95% CI [0.79, 4.18]. Thus, the effect of GMA on both CWB-I and CWB-O was negative at low levels of agreeableness. However, the effect of GMA on CWB-I was also negative at moderately high levels of agreeableness and non-significant beyond 1.21 SD above mean agreeableness.

Emotional Stability

In Sample 1, there was a significant interactive influence of emotional stability on CWB-I ($b = 0.06$, $p = 0.003$) and CWB-O ($b = 0.07$, $p = 0.003$). However, when quadratic controls were included in the model (Cortina, 1993), the interactive influence on CWB-I was no longer significant and, therefore, the JN 95% CI and corresponding simple slopes tests are not reported. Emotional stability accounted 95.7% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.059$). The effect of GMA on CWB-O was significant outside the JN 95% CI [-1.43, 1.78] such that the effect was negative at very low levels of stability and positive at very high levels of stability.

In Sample 2, emotional stability and GMA also showed a significant interactive effect on CWB-I ($b = 0.15$, $p < 0.001$) and CWB-O ($b = 0.12$, $p < 0.001$). Emotional stability accounted for 30.6% of the variability in the relationship between GMA and CWB-I ($R^2 = 0.151$) and 33.1% of the variability in the relationship between GMA and CWB-O ($R^2 = 0.172$). The effect of GMA on CWB-I was significant outside the JN 95% CI [0.89, 2.23] and the effect of GMA on CWB-O was significant outside the JN 95% CI [0.68, 3.00] such that the effect of GMA on both CWB-I and CWB-O was significant at low levels of emotional stability.

Summary

Overall, in Sample 1, results for CWB-I were consistent with the form hypothesized for conscientiousness and emotional stability. However, results for CWB-O were not consistent with the form hypothesized in Sample 1. Although results suggest a negative effect of GMA on CWB-O at very low emotional stability, results also show a *positive* effect of GMA on CWB-O at high levels of agreeableness and very high levels of emotional stability. Importantly, these differences in subdimension and trait-level results were not replicated in Sample 2. Rather, results in Sample 2 were generally consistent across all three traits and both CWB subdimensions. Thus, Sample 1 differences reported here should be interpreted with caution.

Table 13 Results of regression analyses for CWB-I and CWB-O predicted by lower-order traits and GMA in Sample 1

Variable	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Intercept	0.00	0.03	0.926	0.00	0.02	0.949
GMA	-0.07	0.03	0.012	-0.03	0.03	0.384
Consc	-0.22	0.03	< 0.001	-0.39	0.03	< 0.001
Consc. x GMA	0.07	0.03	0.014	0.04	0.03	0.093
R^2 (ΔR^2)	0.058 (0.005)			0.158 (0.002)		
ΔR^2_{mo}	0.412			0.676		
Agreeableness						
Intercept	0.00	0.03	0.891	-0.01	0.03	0.783
GMA	-0.03	0.03	0.236	0.02	0.03	0.401
Agr	-0.36	0.03	< 0.001	-0.29	0.03	< 0.001
Agr. x GMA	0.04	0.02	0.145	0.07	0.03	0.008
R^2 (ΔR^2)	0.134 (0.001)			0.090 (0.006)		
ΔR^2_{mo}	0.501			0.974		
Emotional stability						
Intercept	0.00	0.03	0.999	0.00	0.03	0.998
GMA	-0.07	0.03	0.019	-0.01	0.03	0.732
ES	-0.15	0.03	< 0.001	-0.22	0.03	< 0.001
ES x GMA	0.06[a]	0.03	0.039	0.07	0.03	0.006
R^2 (ΔR^2)	0.032 (0.003)			0.059 (0.006)		
ΔR^2_{mo}	0.384			0.957		

Models were also run in which quadratic terms were included for all predictors per Cortina (1993). [a]When quadratic terms were included, the GMA-emotional stability interaction was no longer significant; inclusion of quadratic terms did not change the significance of any other interaction terms. P-values are two-tailed. $N = 1{,}412$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *Consc*. Conscientiousness; *Agr*. Agreeableness; *ES* Emotional stability

Table 14 Results of simple slope tests for the effect of GMA on CWB-I and CWB-O at low, average, and high levels of lower-order traits in Sample 1

	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Low (-1 SD)	-0.14	0.03	< 0.001			
Average	-0.07	0.03	0.012			
High (+ 1 SD)	-0.01	0.04	1.000			
95% CI	[0.12, 13.34]					
Agreeableness						
Low (-1 SD)				-0.05	0.04	0.554
Average				0.02	0.03	0.402
High (+ 1 SD)				0.10	0.03	0.012
95% CI				[-4.17, 0.59]		
Emotional stability						
Low (-1 SD)				-0.08	0.04	0.100
Average				-0.01	0.03	0.732
High (+ 1 SD)				0.06	0.04	0.202
95% CI				[-1.43, 1.78]		

Results not reported for non-significant interactions (see Table 13). 95% confidence interval applies JN technique to continuous variables (see Bauer & Curran, 2005). When the respective personality trait is outside the confidence interval, the slope of general mental ability (GMA) is significant. Low indicates 1 standard deviation below the mean; high indicates one standard deviation above the mean. JN alpha and simple slopes p-values are Bonferroni corrected (Bauer & Curran, 2005). P-values are two-tailed. $N = 1{,}412$; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB

Table 15 Results of regression analyses for CWB-I and CWB-O predicted by lower-order traits and GMA in Sample 2

Variable	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Intercept	-0.02	0.03	0.402	-0.01	0.02	0.558
GMA	-0.17	0.02	< 0.001	-0.11	0.02	< 0.001
Consc	-0.41	0.02	< 0.001	-0.50	0.02	< 0.001
Consc. x GMA	0.15	0.02	< 0.001	0.10	0.03	< 0.001
R^2 (ΔR^2)	0.258 (0.021)			0.307 (0.009)		
ΔR^2_{mo}	0.380			0.377		
Agreeableness						
Intercept	-0.01	0.03	0.788	-0.01	0.03	0.787
GMA	-0.22	0.03	< 0.001	-0.17	0.03	< 0.001
Agr	-0.27	0.03	< 0.001	-0.26	0.03	< 0.001
Agr. x GMA	0.11	0.03	< 0.001	0.11	0.03	< 0.001
R^2 (ΔR^2)	0.135 (0.011)			0.108 (0.011)		
ΔR^2_{mo}	0.171			0.257		
Emotional Stability						
Intercept	-0.01	0.03	0.656	-0.01	0.03	0.719
GMA	-0.21	0.03	< 0.001	-0.16	0.03	< 0.001
ES	-0.27	0.03	< 0.001	-0.35	0.03	< 0.001
ES x GMA	0.15	0.03	< 0.001	0.12	0.03	< 0.001
R^2 (ΔR^2)	0.151 (0.022)			0.172 (0.014)		
ΔR^2_{mo}	0.306			0.331		

Models were also run in which quadratic terms were included for all predictors per Cortina (1993). In no case did inclusion of the quadratic term change the significance of the interaction term. $N = 1{,}241$; *GMA* General mental ability; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB; *Consc*. Conscientiousness; *Agr*. Aagreeableness; *ES* Emotional stability

Table 16 Results of simple slope tests for the effect of GMA on CWB-I and CWB-O at low, average, and high levels of lower-order traits in Sample 2

	CWB-I			CWB-O		
	b	SE	*p*	*b*	SE	*p*
Conscientiousness						
Low (-1 SD)	-0.32	0.04	< 0.001	-0.21	0.04	< 0.001
Average	-0.17	0.02	< 0.001	-0.11	0.02	< 0.001
High (+1 SD)	-0.02	0.03	1.000	-0.01	0.03	1.000
95% CI	[0.71, 1.74]			[0.52, 2.68]		
Agreeableness						
Low (-1 SD)	-0.33	0.04	< 0.001	-0.29	0.05	< 0.001
Average	-0.22	0.03	< 0.001	-0.17	0.03	< 0.001
High (+1 SD)	-0.11	0.03	0.002	-0.06	0.04	0.314
95% CI	[1.21, 4.26]			[0.79, 4.18]		
Emotional Stability						
Low (-1 SD)	-0.36	0.04	< 0.001	-0.28	0.04	< 0.001
Average	-0.21	0.03	< 0.001	-0.16	0.03	< 0.001
High (+1 SD)	-0.05	0.03	0.216	-0.03	0.04	0.794
95% CI	[0.89, 2.23]			[0.68, 3.00]		

95% confidence interval applies JN technique to continuous variables (see Bauer & Curran, 2005). When the respective personality trait is outside the confidence interval, the slope of general mental ability (GMA) is significant. Low indicates 1 standard deviation below the mean; high indicates one standard deviation above the mean. JN alpha and simple slopes p-values are Bonferroni corrected (Bauer & Curran, 2005). *P*-values are two-tailed. $N = 1{,}241$; *CWB* Counterproductive work behavior; *CWB-I* Interpersonal CWB; *CWB-O* Organizational CWB

Appendix 2

Simulated Selection Analyses

To evaluate the practical impact of results, we conducted simulation analyses in which we compared CWB among "applicants" selected using a stability-GMA interaction model and model with only stability included as a predictor. We compared the interaction model to a stability-only model because stability-related traits (i.e., conscientiousness, agreeableness, and emotional stability) have been shown to negatively predict CWB and are used in selection systems. In contrast, prior research does not consistently support a GMA-CWB relationship, and we are not aware of any selection systems that use GMA to predict CWB.

First, we simulated applicant pools by randomly selecting people from our existing samples. We selected 1,000 samples of 100 people each from each of the samples, as well as from the combined samples. Next, we used coefficient estimates from the stability-only and interaction model to predict CWB and identified applicants to "screen out" (i.e., highest 10% and 20% in CWB). Notably, we cross-validated parameters by using coefficients estimated from each sample to predict CWB in both samples separately, as well as the combined samples. We then compared the observed average and maximum CWB scores of applicants that were identified by the interaction versus stability-only model. A model was considered to have "won" if observed CWB was more than 0.10 SD above that predicted by the other model. The number of "wins" for each model are shown in Tables 17 and 18 for average and maximum CWB, respectively. Results show that the interaction model was better than the stability-only model at predicting high CWB in 10 of 12 comparisons for average CWB and all (12 of 12) comparisons for maximum CWB. That is, the interaction model more accurately "screened out" applicants with the highest CWB in nearly all (22 of 24) comparisons.

Table 17 Comparison of interaction and stability-only model "wins" in simulated selection comparison for average CWB

Simulated applicant sample pool	Using coefficients estimated in Sample 1: % interaction wins (% stability wins)		Using Coefficients estimated in Sample 2: % interaction wins (% stability Wins)	
	Highest 10% in CWB	Highest 20% in CWB	Highest 10% in CWB	Highest 20% in CWB
Sample 1	**47% (23%)**	**20% (16%)**	38% (40%)	20% (38%)
Sample 2	**86% (2%)**	**78% (3%)**	**97% (0%)**	**80% (6%)**
Combined samples	**66% (10%)**	**56% (6%)**	**79% (7%)**	**54% (14%)**

A model was determined to have "won" if the observed maximum CWB for its identified applicants was 0.10 SD more than the observed maximum CWB of the alternate model. Number of "ties" can be calculated by adding the number of interaction model wins (i.e., number outside parentheses) and the number of stability-only model wins (i.e., the number inside the parentheses), and subtracting the total from 100%. Boldface indicates the interaction model outperformed the stability-only model

Table 18 Comparison of interaction and stability-only model "wins" in simulated selection comparison for maximum CWB

Simulated applicant sample pool	Using coefficients estimated in Sample 1: % interaction wins (% stability wins)		Using coefficients estimated in Sample 2: % interaction wins (% stability wins)	
	Highest 10% in CWB	Highest 20% in CWB	Highest 10% in CWB	Highest 20% in CWB
Sample 1	**31% (4%)**	**12% (3%)**	**37% (10%)**	**19% (7%)**
Sample 2	**78% (2%)**	**49% (3%)**	**96% (1%)**	**51% (29%)**
Combined samples	**42% (5%)**	**25% (3%)**	**58% (7%)**	**31% (6%)**

A model was determined to have "won" if the observed maximum CWB for its identified applicants was 0.10 SD more than the observed maximum CWB of the alternate model. Number of "ties" can be calculated by adding the number of interaction model wins (i.e., number outside parentheses) and the number of stability-only model wins (i.e., the number inside the parentheses), and subtracting the total from 100. Boldface indicates the interaction model outperformed the stability-only model

Appendix 3

Latent Profile Analyses

As one reviewer noted, another potential way of examining the veracity of a moderation hypothesis is to estimate latent profiles among the variables of interest. Therefore, we estimated a latent profile model using the 'tidyLPA' package (Rosenberg et al., 2018) in R. We utilized the IRT-derived scores for each of the three lower-order stability traits (i.e., conscientiousness, agreeableness, and emotional stability), along with the two subdimensions of CWB (CWB-I and CWB-O). To rule out any effects driven by the lower-order plasticity traits (i.e., extraversion and openness), we also included IRT scores for these two FFM dimensions.

Considering the Bayesian Information Criterion (BIC; smallest value criteria), Entropy (highest value criteria), smallest class size (in proportions, all classes equal to or greater than 0.10), and overall interpretability, we concluded that in both samples a 4-class solution was best. Statistics for model evaluation are shown in Table 19. As shown in Figure 7, the class with the highest level of CWBs (i.e., Class 3 in Sample 1, and Class 2 in Sample 2) also showed the lowest levels of conscientiousness, agreeableness, and emotional stability,[6] as well as the lowest GMA compared to other classes. Thus, individuals with very high CWB (11% of Sample 1; and 19% of Sample 2) were also low in *both* stability and GMA. The class with the lowest level of CWB (i.e., Class 2 in Sample 1, and Class 1 in Sample 2) had high levels of stability but average levels of GMA. That is, GMA was not a distinguishing feature for those in classes with low CWB (20% in both samples), whereas these classes were high in stability. These results are fully consistent with our hypothesized interaction; individuals with low GMA *and* low stability have high CWB while individuals with high stability (but average GMA) have low CWB. Individuals with levels of CWB closer to average (i.e., the remaining classes: Classes 1 and 4 in Sample 1; Classes 3 and 4 in Sample 2) generally showed average levels of other traits. Notably, there were no clear systematic differences in openness and extraversion across classes and samples. Thus, we believe that LPA shows the same effects that are more explicitly tested in our moderated multiple regression analysis.

Table 19 Fit statistics for latent profile model evaluations

# of Classes	Sample 1			Sample 2		
	BIC	Entropy	Smallest class size	BIC	Entropy	Smallest class size
1	32,165	1.00	1.00	28,280	1.00	1.00
2	31,135	0.68	0.36	26,949	0.81	0.29
3	30,651	0.75	0.17	26,021	0.87	0.22
4	**30,446**	**0.77**	**0.11**	**25,324**	**0.90**	**0.11**
5	30,401	0.68	0.11	25,083	0.90	0.07
6	30,335	0.69	0.08	24,949	0.87	0.06
7	30,249	0.75	0.02	24,696	0.89	0.07

[6] The only exception was that in Sample 2, Class 4 showed lower Agreeableness than Class 2.

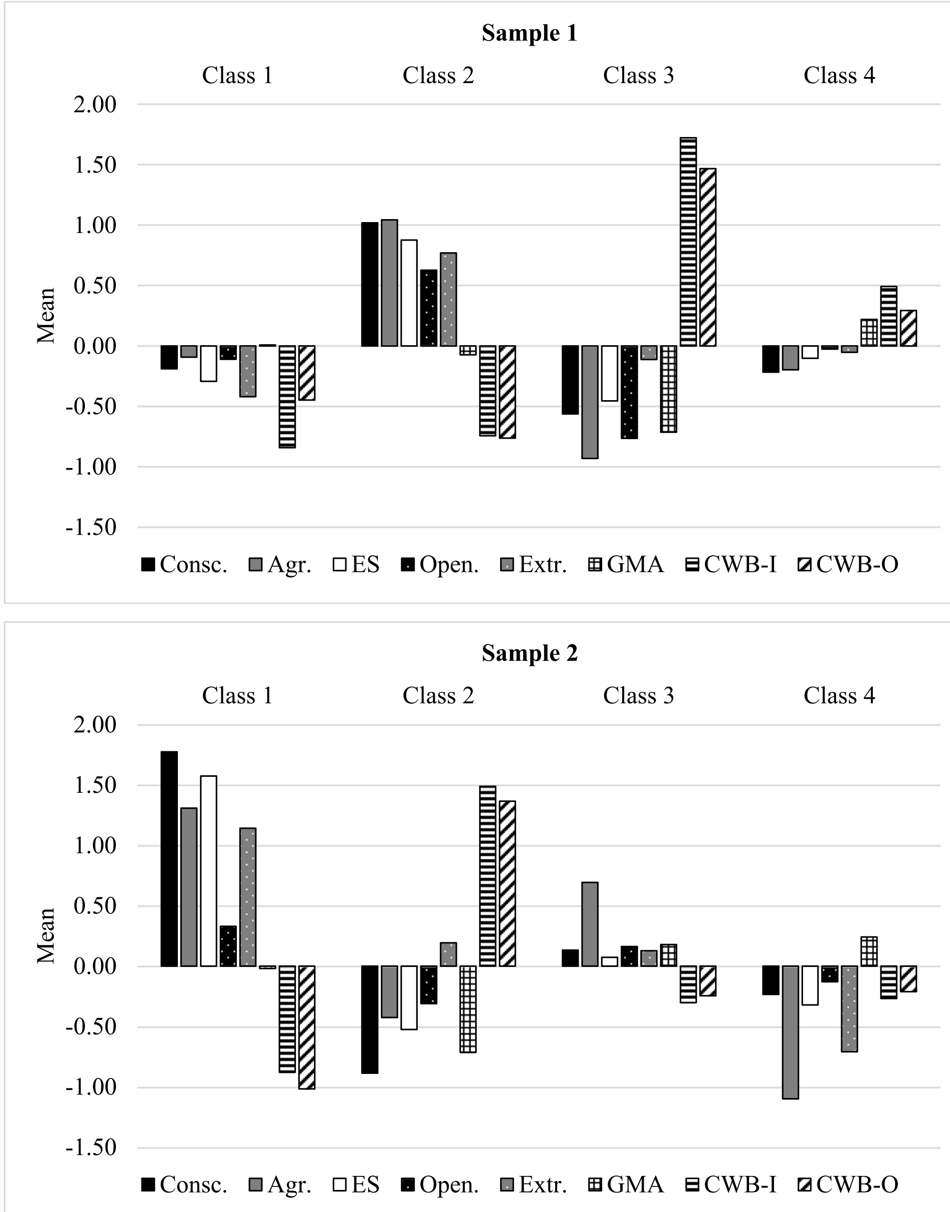

Fig. 7 Mean plots for each in latent profile class in Sample 1 (top) and Sample 2 (bottom)

Data Transparency Appendix

Sample 1 data reported in this manuscript have been previously published. Findings from the data collection have been reported in separate manuscripts. MS 1 (published) focuses on general mental ability (GMA), openness to experience, and creative achievement. MS 2 (published) focuses on GMA only and gender of participants. MS 2 (current manuscript) focuses on GMA, the Five Factor Model traits comprising meta-trait stability (conscientiousness, agreeableness, and emotional stability), and counterproductive work behavior. The table below displays which data variables appear in each study, as well as the current status of each study. Sample 2 data reported in this manuscript have not been previously published.

Table 20 Table of manuscripts connected to data collection

Variables in the Complete Dataset	MS 1 (published)	MS 2 (published)	MS 3 (current MS)
General mental ability	X	X	X
Openness to experience	X		
Conscientiousness			X
Agreeableness			X
Emotional stability			X
Extraversion			
Creative achievement	X		
Job satisfaction			
Need for closure			
Counterproductive work behavior			X
Organizational citizenship behavior			
Demographics (e.g., gender, race, age)		X	

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

We have no conflicts of interest to disclose. This article is based on the dissertation completed by Harris (2020). The work of Alexandra M. Harris-Watson was supported in part by the National Science Foundation Research Fellowship Program (DGE-1443117), and the work of Nathan T. Carter was supported in part by the National Science Foundation (SES1561070). Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the view of the National Science Foundation.

References

Ackerman, P., & Heggestad, E. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*, 219–245. https://doi.org/10.1037//0033-2909.121.2.219

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology, 90*(1), 94–107. https://doi.org/10.1037/0021-9010.90.1.94

Allemand, M., Job, V., & Mroczek, D. K. (2019). Self-control development in adolescence predicts love and work in adulthood. *Journal of Personality and Social Psychology, 117*(3), 621–634. https://doi.org/10.1037/pspp0000229

Ayduk, O., Rodriguez, M. L., Mischel, W., Shoda, Y., & Wright, J. (2007). Verbal intelligence and self-regulatory competencies: Joint predictors of boys' aggression. *Journal of Research in Personality, 41*(2), 374–388. https://doi.org/10.1016/j.jrp.2006.04.008

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40*(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5

Barrick, M. R., & Mount, M. K. (1991). The Big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the New Millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*(1–2), 9–30. https://doi.org/10.1111/1468-2389.00160

Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology, 85*(3), 349–360. https://doi.org/10.1037/0021-9010.85.3.349

Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology, 97*(3), 613–636. https://doi.org/10.1037/a0026739

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology, 92*(2), 410–424. https://doi.org/10.1037/0021-9010.92.2.410

Borman, W., & Motowidlo, S. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98).

Bowling, N. A., & Eschleman, K. J. (2010). Employee personality as a moderator of the relationships between work stressors and counterproductive work behavior. *Journal of Occupational Health Psychology, 15*(1), 91–103. https://doi.org/10.1037/a0017326

Bowling, N. A., & Gruys, M. L. (2010). Overlooked issues in the conceptualization and measurement of counterproductive work behavior. *Human Resource Management Review, 20*(1), 54–61. https://doi.org/10.1016/j.hrmr.2009.03.008

Busenbark, J. R., Graffin, S. D., Campbell, R. J., & Lee, E. Y. (2022). A Marginal Effects Approach to Interpreting Main Effects and Moderation. *Organizational Research Methods, 25*(1), 147–169. https://doi.org/10.1177/1094428120976838

Cao, M., Song, Q. C., & Tay, L. (2018). Detecting curvilinear relationships: A comparison of scoring approaches based on different item response models. *International Journal of Testing, 18*(2), 178–205. https://doi.org/10.1080/15305058.2017.1345913

Carter, N. T., Dalal, D. K., Boyce, A. S., O'Connell, M. S., Kung, M.-C., & Delgado, K. M. (2014). Uncovering curvilinear relationships between conscientiousness and job performance: How theoretically appropriate measurement makes an empirical difference. *Journal of Applied Psychology, 99*, 564–586. https://doi.org/10.1037/a0034688

Carter, N. T., Dalal, D. K., Guan, L., LoPilato, A. C., & Withrow, S. A. (2017). Item response theory scoring and the detection of curvilinear relationships. *Psychological Methods, 22*, 191–203. https://doi.org/10.1037/met0000101

Carver, C. S., & Miller, C. J. (2006). Relations of serotonin function to personality: Current views and a key methodological issue. *Psychiatry Research, 144*(1), 1–15. https://doi.org/10.1016/j.psychres.2006.03.013

Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research, 66*(9), 1261–1266. https://doi.org/10.1016/j.jbusres.2012.02.023

Choi, J. Y., Miao, C., Oh, I. S., Berry, C. M., & Kim, K. (2018). Relative importance of major job performance dimensions in determining supervisors' overall job performance ratings. *Canadian Journal of Administrative Sciences, 36*(3), 377–389. https://doi.org/10.1002/CJAS.1495

Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management, 19*(4), 915–922. https://doi.org/10.1177/014920639301900411

Costa, P. T., & McCrae, R. R. (1992). Professional manual: Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI). *Psychological Assessment Resources*.

Cullen, M. J., & Sackett, P. R. (2004). Personality and counterproductive workplace behavior. In M. Barrick & A. M. Ryan (Eds.), *Personality and Work: Reconsidering the Role of Personality in Organizations*. John Wiley & Sons.

de Vries, R. E., & van Gelder, J.-L. (2013). Tales of two self-control scales: Relations with Five-Factor and HEXACO traits. *Personality and Individual Differences, 54*(6), 756–760. https://doi.org/10.1016/j.paid.2012.12.023

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*(6), 1138–1151. https://doi.org/10.1037/0022-3514.91.6.1138

DeYoung, C. G. (2015). Cybernetic Big Five Theory. *Journal of Research in Personality, 56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004

DeYoung, C. G., Hirsch, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the Big Five. *Psychological Science, 21*(6), 820–828. https://doi.org/10.1177/0956797610370159

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*(4), 533–552. https://doi.org/10.1016/S0191-8869(01)00171-4

DeYoung, C. G., Peterson, J. B., Séguin, J. R., & Tremblay, R. E. (2008). Externalizing behavior and the higher order factors of the Big Five. *Journal of Abnormal Psychology (1965), 117*(4), 947–953. https://doi.org/10.1037/a0013742

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*(6), 1246–1256. https://doi.org/10.1037/0022-3514.73.6.1246

Dilchert, S., Ones, D. S., Davis, R. D., & Rostow, C. D. (2007). Cognitive ability predicts objectively measured counterproductive work behaviors. *Journal of Applied Psychology, 92*(3), 616–627. https://doi.org/10.1037/0021-9010.92.3.616

Dunlop, P. D., & Lee, K. (2004). Workplace deviance, organizational citizenship behavior, and business unit performance: The bad apples do spoil the whole barrel. *Journal of Organizational Behavior, 25*(1), 67–80. https://doi.org/10.1002/job.243

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology, 61*(3), 579–616. https://doi.org/10.1111/j.1744-6570.2008.00123.x

Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive work behavior (CWB) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of Vocational Behavior, 59*(3), 291–309. https://doi.org/10.1006/jvbe.2001.1803

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Gonzalez-Mulé, E., Mount, M. K., & Oh, I.-S. (2014). A meta-analysis of the relationship between general mental ability and nontask performance. *Journal of Applied Psychology, 99*(6), 1222–1243. https://doi.org/10.1037/a0037547

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.

Guay, R. P., Choi, D., Oh, I.-S., Mitchell, M. S., Mount, M. K., & Shin, K.-H. (2016). Why people harm the organization and its members: Relationships among personality, organizational commitment, and workplace deviance. *Human Performance, 29*(1), 1–15. https://doi.org/10.1080/08959285.2015.1120305

Harris, A. (2020). *The inhibitory effect of general mental ability on the link between personality and counterproductive work behavior*. [Doctoral Dissertation, University of Georgia].

Harris, A., McMillan, J. T., Listyg, B. J., Matzen, L. E., & Carter, N. T. (2020). Measuring intelligence with the Sandia Matrices: Psychometric review and recommendations for free Raven-like item sets. *Personnel Assessment and Decisions*, *6*(3). https://doi.org/10.25035/pad.2020.03.006

Hirschi, T., & Gottfredson, M. R. (1994). The generality of deviance. In T. Hirschi & M. R. Gottfredson (Eds.), The generality of deviance (pp. 1–22). New Brunswick, NJ: Transaction Publishers.

Hirsh, J. B., DeYoung, C. G., & Peterson, J. B. (2009). Metatraits of the Big Five Differentially Predict Engagement and Restraint of Behavior. *Journal of Personality, 77*(4), 1085–1102. https://doi.org/10.1111/j.1467-6494.2009.00575.x

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9*(1–2), 152–194. https://doi.org/10.1111/1468-2389.00171

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Krischer, M. M., Penney, L. M., & Hunter, E. M. (2010). Can counterproductive work behaviors be productive? CWB as emotion-focused coping. *Journal of Occupational Health Psychology, 15*(2), 154–166. https://doi.org/10.1037/a0018349

LaPalme, M., Tay, L., & Wang, W. (2018). A within-person examination of the ideal-point response process. *Psychological Assessment, 30*(5), 567–581. https://doi.org/10.1037/pas0000499

Lee, Y.-S., & Luna-Bazaldua, D. A. (2019). How to conduct a study with diagnostic models. In M. von Davier & Y.-S. Lee (Eds.), *Handbook of Diagnostic Classification Models: Models and Model Extensions, Applications, Software Packages* (pp. 525–545). Springer International Publishing. https://doi.org/10.1007/978-3-030-05584-4

Lievens, F., Conway, J. M., & De Corte, W. (2008). The relative importance of task, citizenship and counterproductive performance to job performance ratings: Do rater source and team-based culture matter? *Journal of Occupational and Organizational Psychology, 81*(1), 11–27. https://doi.org/10.1348/096317907X182971

Liu, H., & Yuan, K.-H. (2020). New measures of effect size in moderation analysis. *Psychological Methods*. https://doi.org/10.1037/met0000371

Lykken, D. (1991). What's wrong with psychology anyway? In D. Cicchetti & W. M. Grove (Eds.), Thinking clearly about psychology, Vol. 1: Matters of public interest. (pp. 3–39). Minneapolis: University of Minnesota Press.

Lynam, D., Moffitt, T., & Stouthamer-Loeber, M. (1993). Explaining the relation between IQ and delinquency: Class, race, test motivation, school failure, or self-control? *Journal of Abnormal Psychology (1965), 102*(2), 187–196. https://doi.org/10.1037/0021-843X.102.2.187

Maples, J. L., Guan, L., Carter, N. T., & Miller, J. D. (2014). A test of the International Personality Item Pool representation of the Revised NEO Personality Inventory and development of a 120-Item IPIP-based measure of the Five-Factor Model. *Psychological Assessment, 26*(4), 1070–1084. https://doi.org/10.1037/pas0000004

Marcus, B. (2003). An empirical examination of the construct validity of two alternative self-control measures. *Educational and Psychological Measurement, 63*(4), 674–706. https://doi.org/10.1177/0013164403251329

Marcus, B., & Schuler, H. (2004). Antecedents of counterproductive behavior at work: A general perspective. *Journal of Applied Psychology, 89*(4), 647–660. https://doi.org/10.1037/0021-9010.89.4.647

Marcus, B., Wagner, U., Poole, A., Powell, D. M., & Carswell, J. (2009). The relationship of GMA to counterproductive work behavior revisited. *European Journal of Personality, 23*(6), 489–507. https://doi.org/10.1002/per.728

Matzen, L. E., Benz, Z. O., Dixon, K. R., Posey, J., Kroger, J. K., & Speed, A. E. (2010). Recreating Raven's: Software for systematically generating large numbers of Raven-like matrix problems with normed properties. *Behavior Research Methods, 42*(2), 525–541. https://doi.org/10.3758/BRM.42.2.525

Maydeu-Olivares, A., & Joe, H. (2014). Assessing approximate fit in categorical data analysis. *Multivariate Behavioral Research, 49*(4), 305–328. https://doi.org/10.1080/00273171.2014.911075

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43*(2), 335–354. https://doi.org/10.1111/j.1744-6570.1990.tb01562.x

Meier, L. L., & Spector, P. E. (2013). Reciprocal effects of work stressors and counterproductive work behavior: A five-wave longitudinal study. *Journal of Applied Psychology, 98*(3), 529–539. https://doi.org/10.1037/a0031732

Moffitt, T. E., & Silva, P. A. (1988). IQ and delinquency: A direct test of the differential detection hypothesis. *Journal of Abnormal Psychology (1965), 97*(3), 330–333. https://doi.org/10.1037/0021-843X.97.3.330

Monroe, S., & Cai, L. (2015). Evaluating structural equation models for categorical outcomes: A new test statistic and a practical challenge of interpretation. *Multivariate Behavioral Research, 50*(6), 569–583. https://doi.org/10.1080/00273171.2015.1032398

Morse, B. J., Johanson, G. A., & Griffeth, R. W. (2012). Using the graded response model to control spurious interactions in moderated multiple regression. *Applied Psychological Measurement, 36*(2), 122–146. https://doi.org/10.1177/0146621612438725

Mount, M., Ilies, R., & Johnson, E. (2006). Relationship of personality traits and counterproductive work behaviors: The mediating effects of job satisfaction. *Personnel Psychology, 59*(3), 591–622. https://doi.org/10.1111/j.1744-6570.2006.00048.x

O'Gorman, J. G., & Baxter, E. (2002). Self-control as a personality measure. *Personality and Individual Differences, 32*(3), 533–539. https://doi.org/10.1016/S0191-8869(01)00055-1

Olson, K. R. (2005). Engagement and Self-Control: Superordinate dimensions of Big Five traits. *Personality and Individual Differences, 38*(7), 1689–1700. https://doi.org/10.1016/j.paid.2004.11.003

Ones, D. (2002). Introduction to the Special Issue on Counterproductive Behaviors at Work. *International Journal of Selection and Assessment, 10*. https://doi.org/10.1111/1468-2389.00188

Ones, D. S., Dilchert, S., & Viswesvaran, C. (2012). Cognitive abilities. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 179–224). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199732579.001.0001

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609–626. https://doi.org/10.1002/(SICI)1099-1379(199611)17:6%3c609::AID-JOB1828%3e3.0.CO;2-K

Oppler, S. H., McCloy, R. A., Peterson, N. G., Russell, T. L., & Campbell, J. P. (2001). The prediction of multiple components of entry-level performance. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the Limits in Personnel Selection and Classification*. Lawrence Erlbaum Associates Publishers.

Penney, L. M., & Spector, P. E. (2005). Job stress, incivility, and counterproductive work behavior (CWB): The moderating role of negative affectivity. *Journal of Organizational Behavior, 26*(7), 777–796. https://doi.org/10.1002/job.336

Postlethwaite, B., Robbins, S., Rickerson, J., & McKinniss, T. (2009). The moderation of conscientiousness by cognitive ability when predicting workplace safety behavior. *Personality and Individual Differences, 47*(7), 711–716. https://doi.org/10.1016/j.paid.2009.06.008

Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology, 38*(3), 931–964. https://doi.org/10.1111/j.1745-9125.2000.tb00911.x

Raven, J. C., Court, J. H., & Raven, J. (1998). Raven's progressive matrices*.* Oxford Psychologists Press.

Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement, 24*(1), 3–32. https://doi.org/10.1177/01466216000241001

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(1), 1–36. https://doi.org/10.18637/jss.v048.i02

Rosenberg, et al. (2018). TidyLPA: An r package to easily carry out latent profile analysis (LPA) using open-source or commercial software. *Journal of Open Source Software, 3*(30), 978. https://doi.org/10.21105/joss.00978

Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology, 87*(1), 66–80. https://doi.org/10.1037//0021-9010.87.1.66

Sackett, P. R., Berry, C. M., Wiemann, S. A., & Laczo, R. M. (2006). Citizenship and counterproductive behavior: Clarifying relations between the two domains. *Human Performance, 19*(4), 441–464. https://doi.org/10.1207/s15327043hup1904_7

Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*(1), 419–450. https://doi.org/10.1146/annurev.psych.59.103006.093716

Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). Springer. https://doi.org/10.1007/978-1-4757-2691-6_5

Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L., Church, A. T., Katigbak, M. S., Somer, O., Szarota, P., Szirmák, Z., & Zhou, X. (2014). A basic bivariate structure of personality attributes evident across nine languages. *Journal of Personality, 82*(1), 1–14. https://doi.org/10.1111/jopy.12028

Spector, P. E. (2011). The relationship of personality to counterproductive work behavior (CWB): An integration of perspectives. *Human Resource Management Review, 21*(4), 342–352. https://doi.org/10.1016/j.hrmr.2010.10.002

Spector, P. E., & Fox, S. (2002). An emotion-centered model of voluntary work behavior: Some parallels between counterproductive work behavior and organizational citizenship behavior. *Human Resource Management Review, 12*(2), 269–292. https://doi.org/10.1016/S1053-4822(02)00049-9

Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior, 68*(3), 446–460. https://doi.org/10.1016/j.jvb.2005.10.005

The International Cognitive Ability Resource Team. (2014). https://icar-project.com/

van Zyl, C. J. J., & De Bruin, G. P. (2018). Predicting counterproductive work behavior with narrow personality traits: A nuanced examination using quantile regression. *Personality and Individual Differences, 131*, 45–50. https://doi.org/10.1016/j.paid.2018.04.014

Vazsonyi, A. T., Pickering, L. E., Junger, M., & Hessing, D. (2001). An empirical test of a general theory of crime: A four-nation comparative study of self-control and the prediction of deviance. *The Journal of Research in Crime and Delinquency, 38*(2), 91–131. https://doi.org/10.1177/0022427801038002001